[redacted]

NASA

Technical Memorandum 80250

[redacted]

PROCEEDINGS OF THE TENTH ANNUAL PRECISE TIME AND TIME INTERVAL (PTTI) APPLICATIONS AND PLANNING MEETING

NOVEMBER 1978

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

For information concerning availability of this document, contact:

TID
Code 250
Goddard Space Flight Center

Tel. 301-344-8488

TM 80250

PROCEEDINGS
OF THE TENTH ANNUAL
PRECISE TIME AND TIME INTERVAL
(PTTI)
APPLICATIONS AND PLANNING MEETING

Held at the Naval Research Laboratory

November 28–30, 1978

Sponsored by

Naval Electronic Systems Command
NASA Goddard Space Flight Center
Naval Research Laboratory
Naval Observatory
Defense Communications Agency

Prepared by
GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland 20771

Page intentionally left blank

EXECUTIVE COMMITTEE

Ralph T. Allen
Naval Electronic Systems Command

Charles A. Bartholomew
Naval Research Laboratory

Andrew R. Chi
NASA Goddard Space Flight Center

Laura C. Fisher
Naval Observatory

Andrew C. Johnson
Naval Observatory

James A. Murray, Jr.
Naval Research Laboratory

Dr. Harris A. Stover
Defense Communications Agency

Schuyler C. Wardrip
NASA Goddard Space Flight Center

GENERAL CHAIRMAN

DR. GART WESTERHOUT
Naval Observatory

TECHNICAL PROGRAM COMMITTEE

DAVID W. ALLAN - CHAIRMAN
National Bureau of Standards

DR. THOMAS E. CLARK, JR.
NASA Goddard Space Flight Center

DR. HELMUT W. HELLWIG
National Bureau of Standards

DR. CHARLES F. MARTIN
Defense Mapping Agency

DR. HARRIS A. STOVER
Defense Communications Agency

SCHUYLER C. WARDRIP
NASA Goddard Space Flight Center

DR. GERNOT M. R. WINKLER
Naval Observatory

EDITORIAL COMMITTEE

L. J. RUEGER - CHAIRMAN
Johns Hopkins University
Applied Physics Laboratory

RALPH T. ALLEN
Naval Electronic Systems Command

ERIC BLOMBERG
Smithsonian Astrophysical Observatory

JAMES A. BUISSON
Naval Research Laboratory

SANDRA HOWE
National Bureau of Standards

KENNITH PUTKOVICH
Naval Observatory

DR. HARRIS A. STOVER
Defense Communications Agency

FRANCIS MOORE – SECRETARY
Applied Physics Laboratory

SESSION CHAIRMEN

SESSION I
Dr. Gernot M. R. Winkler
Naval Observatory

SESSION II
Dr. Helmut W. Hellwig
National Bureau of Standards

SESSION III
David W. Allan
National Bureau of Standards

SESSION IV
Dr. Harris A. Stover
Defense Communications Agency

SESSION V
Dr. Charles F. Martin
Defense Mapping Agency

SESSION VI
Dr. Thomas E. Clark, Jr.
NASA Goddard Space Flight Center

ARRANGEMENTS

James A. Murray, Jr., NRL
Charles Bartholomew, NRL
Herbert S. Poole, NRL

FINANCE COMMITTEE

Laura C. Fisher, USNO
Schuyler C. Wardrip, GSFC

PUBLICATIONS

Lois Schrom, GSFC
Schuyler C. Wardrip, GSFC

TECHNICAL ASSISTANCE

David Phillips, NRL
William B. Nafey, NRL

PRINTING

Charles V. Hardesty, GSFC
Donald E. Ellis, GSFC

RECEPTIONISTS

Kathy Booth, NRL
Kathy McDonald, NRL
Marie Nader, GSFC
Dorothy Outlaw, USNO
Emma K. Thomas, GSFC

BANQUET SPEAKER

Dr. R. Grant Athay
High Altitude Observatory, Boulder, Colorado
Subject: Mysteries of the Sun

CALL TO SESSION

Dr. Gart Westerhout
Scientific Director, Naval Observatory

WELCOME ADDRESS

Capt. Edward E. Henifin
Commanding Officer, Naval Research Laboratory

OPENING COMMENTS

Tecwyn Roberts
Director, Networks Directorate
NASA Goddard Space Flight Center

Rear Adm. G. H. Smith
Vice Commander, Naval Electronic Systems Command

Capt. Joseph C. Smith
Superintendent, Naval Observatory

Page intentionally left blank

CONTENTS

SESSION III

SESSION IV

SESSION V

SESSION VI

Page intentionally left blank

FOREWORD

These proceedings contain the papers presented at the Tenth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, including questions and answers following the presentations.

As stated in the Program Introduction, the purpose of the PTTI Applications and Planning Meeting is:

To give PTTI Managers, Systems Engineers, and Program Planners:

- A transparent view of the state-of-the-art
- An opportunity to express needs (current or future)
- A view of important future trends
- A review of relevant past accomplishments.

To provide PTTI users with new and useful applications, procedures, and techniques.

To allow the PTTI researcher to better assess fruitful directions for research efforts.

The 195 registered attendees came from Government, private industry, Universities, and foreign countries.

This year, most sessions were started with invited tutorial papers followed by contributed papers. This format proved very successful; The tutorials set the tone for the following discussions. A successful Discussion Forum on Atomic Frequency Standards was a fitting end to the first day. The tutorials covered the following subjects:

- Composite Oscillator Systems for Meeting User Needs for Time and Frequency
- Quartz Crystal and Super Conductive Resonators and Oscillators
- Time Domain Measurement Systems
- Frequency Domain Measurement Systems
- Clock Performance as a critical Parameter in Navigational Satellite Systems
- Down to Earth Relativity
- Modern Technology for the Determination of UT1 and Polar Motion.

On behalf of the Executive Committee, I want to thank all those who contributed to the success of this year's meeting. Special thanks go to the Chairmen of the Technical Program Committee who carried the major responsibility for this excellent program and to the Session Chairman. Without S. Clark Wardrip, who was on top of absolutely everything, this meeting would have been impossible.

Gart Westerhout

GART WESTERHOUT
General Chairman

Page intentionally left blank

CALL TO SESSION

Dr. Gart Westerhout
U. S. Naval Observatory

DR. WESTERHOUT: Good Morning. I am Gart Westerhout, Scientific Director of the U. S. Naval Observatory. As general chairman for this Tenth Annual Precise Time and Time Interval Applications and Planning Meeting, I call the meeting to session.

This year's meeting features 7 tutorial papers, 20 contributed papers and a discussion forum. The executive committee asked the program committee to draw up a program which emphasized the purpose of the meeting: To provide PTTI managers, engineers, planners, users, and researchers with a review of the state of the art and a discussion of needs and future trends. I recommend that you all read page VI in the program booklet again, so that we remember the reasons for this meeting, which is properly called an Applicatons and Planning Meeting. The executive committee does not want this annual affair to become the equivalent of an IEEE type meeting.

Having said all this, I need to thank our technical program committee and especially its able chairman, David W. Allan, for a job well done. Subjects for the tutorials seem very appropriate and will hopefully contribute to the overall discussion as well as to the purpose of the meeting: to assist PTTI Managers, Systems Engineers and Program Planners in their tasks.

I also wish to thank the members of the executive committee for the smooth manner in which this tenth meeting has come into being. I have spent 9 months trying to find out what I was supposed to do as general chairman, and I'm still looking. Everything happened automatically. That's why I feel it incumbent on me to speak to you for half an hour today; at least that gives me something to do.

Thanks are due especially to S. Clark Wardrip of NASA who handled all the details of putting the meeting together, and to James A. Murray and Charles Bartholomew of NRL who handled the local arrangements.

Before I yield the microphone, two comments: Please use the microphone when you are making comments, and state your name and affiliation. Discussion after each paper is very strongly encouraged and the recording of the discussion is considered to make the proceedings extremely valuable.

And the second comment: Wednesday's banquet speaker, Dr. R.Grant Athay, is a world-renowned solar physicist. I warmly recommend attendance at the banquet for both its tangible and intangible benefits.

I would now like to call on the representatives of some of the agencies who sponsor the PTTI meetings to say a few words of welcome.

WELCOME ADDRESS

Capt. Edward E. Henifin
Commanding Officer
Naval Research Laboratory

On behalf of the Naval Research Laboratory, I would like to welcome you to the Tenth Annual Precise Time and Time Interval Applications and Planning Meeting. It is a pleasure for NRL to act as one of the co-hosts and sponsors of this meeting.

You are about to embark on a series of technical sessions for the next three days to accomplish the goals of the meeting. These sessions are going to be extremely interesting and thought-provoking and, since you all are involved in precise time, I'm sure the sessions will all go according to the planned schedule.

Most individuals are clock-watchers of sorts and from our earliest memories time has played an important role in our lives. From my own experiences, my earliest recollection of time and time measurement goes back to the early school years watching the hands of an old Seth Thomas move ever so slowly toward the three-thirty mark when school would be out. Measurement was the number of days until the next weekend or the next vacation - it was not a precise thing - just a block of days divided into various intervals.

Later on, as I became more aware of the importance of time, I realized that my grandfather and father, both railroad engineers, were extremely interested in exact time, especially that which had to do with trains. On a periodic schedule both had to have their big pocket watches checked by a particular jeweler. The watches had to be precise - that is, would not lose or gain more than a few seconds in a month - I can't remember the number now, but that's probably not important.

Then, there was my first watch - used to check it with the radio daily to make sure it wasn't slow or fast. A minute or two a week made me feel pretty good - I had an accurate watch.

Then, I was exposed to navigation. Time and accuracy took on a new dimension. The daily time ticks with WWV and the chronometer log sun lines; apparent noon; star fixes; etc. required accurate time for accurate position reports. In 1968/69 I encouraged my navigator to write a paper up the chain of command proposing that we replace the chronometers on board ship with the then new Bulova Accutrons - they were advertized as being more accurate/precise than the chronometers we had on board - accurate to tenths of seconds in a 30-day period. As you may know, chronometers had to be overhauled every operating

cycle at costs of about three to four times that of an Accutron. Tenths of seconds or even hundredths of seconds are comprehensible to me but comprehending nano and pico seconds is like trying to visualize a billion dollars all in one-dollar bills - mind boggling.

That's a brief of my experience with time, with the exception of this past year when I met the keeper of the time, Dr. Gernot Winkler, at the Naval Observatory and additionally had the good fortune of being briefed on the PTTI program. I won't say that I understand PTTI, but I do have an appreciation for what it is and what your goals, when achieved, will mean.

NRL is proud of the part it has played in the time and frequency field over the years and is looking forward to a continuing contribution in the future.

It is therefore with great pleasure that on behalf of NRL I again welcome you to this conference.

OPENING COMMENTS

Tecwyn Roberts
Director, Networks Directorate, NASA/Goddard Space Flight Center

As some of you may know, 1978 marks the 20th Anniversary of NASA, and I think it appropriate to reflect a few moments about some of what has been accomplished in those twenty years and where we are going, not only in the precision frequency and time area but also in other areas.

From NASA's inception, the Agency has successfully worked both to advance man's knowledge of the universe, and to use that knowledge to better life on Earth. We are proud of Goddard's role in this effort.

Goddard's early work with communication satellites such as Relay, Syncom and the Applications Technology Satellites (ATS) laid the ground work for low cost transatlantic telephone calls and radio and television broadcast.

The key to future growth in satellite communications is the broadcast satellite which beams signals directly to rooftop size terminals. ATS–6 experiments demonstrated that one satellite in a stationary orbit can serve thousands of users.

Goddard's research with metrology satellites has and is revolutionizing the study and forecasting of global weather conditions. Projects managed by Goddard such as the TIROS series of satellites laid the groundwork for the use of satellites for predicting the weather. TIROS–N, which was launched Oct. 13th, has been active mapping the weather and tracking hurricanes; the first of which was Kendra.

In the early 1960's NASA initiated the Nimbus satellite program which was designed to meet the research needs of atmospheric and earth scientists. The final satellite in this series, Nimbus 7, was launched Oct. 24th in a near polar orbit so as to monitor man made and natural pollutants in the atmosphere which concerns us all.

Goddard is also very much a part of the Global Atmospheric Research Program (GARP) which was established in 1967 under the auspices of the United Nations. The first Global Weather Experiment is scheduled to start in December or January. Data will be collected for one year from over 10,000 sources; platforms on the ground, ships at sea, aircraft and weather satellites operated by the United States, Japan, the European Space Agency (ESA), and the USSR.

NASA's Earth Resources Satellites are providing the means for nations to take constant inventory of the Earth's dwindling supply of natural resources. Goddard's role in this program includes the management of the Landsat satellites. These satellites with special sensors are able to distinguish different crops, measure ground temperature, and determine flood patterns, erosion, and drought areas. Last May NASA launched a satellite to map the day/night temperature differences of the earth's surface. The satellite called the Heat Capacity Mapping Mission (HCMM) will be used with temperature information from Landsat 3 to classify rock formations, monitor soil moisture changes and agriculture stress, and study the effects of urban heat islands on local weather.

Next year NASA will launch its first geological applications satellite termed Magsat. The data Magsat sends back will aid in mineral and oil exploration and in updating world and regional magnetic charts used for navigation and surveying.

Other Goddard projects look beyond the Earth to study the influence of the sun on the Earth and observe the stars and planets to learn more about cosmic processes. In 1979, at the height of the next eleven year cycle of maximum solar activity, NASA will launch Goddard's Solar Maximum Mission (SMM) to study the entire electromagnetic spectrum of solar flares to try and determine their physical origin.

It may interest you to know that it was Goddard's Copernicus satellite that this past June located a second invisible black hole in our galaxy orbiting a giant super star. The black hole is gradually siphoning away the larger star's atmosphere. A gamma-ray observatory scheduled for the mid 1980's will look again at black holes as it maps gamma-ray sources throughout the universe. Studying black holes will extend man's knowledge of physics into aspects of relatively not observable here on Earth.

Through the International Ultraviolet Explorer (IUE) experiment astronomers from NASA, the United Kingdom, and the European Space Agency are continually looking at objects ranging from planets in our solar system to the most distant objects in the universe. Another experiment planned for 1983, the Cosmic Background Explorer Satellite, will use infrared waves to address the question of the very origin of the universe.

Coming to more earthly aspects, in the early 1980's the Space Shuttle will begin making round-trips into space. Eventually, there is expected to be as many as forty missions per year. For Shuttle, Goddard is developing a standardized satellite container called a Multimission Module Spacecraft that can house a diversity of experiments and instruments.

Goddard also has the responsibility for administering the Shuttle Getaway Special Program. Through this program, individuals, industries, research groups, and other nations can place small experiments aboard Shuttle for as little as $3000. In fact, Goddard's Explorer Scout Post is planning a project on Shuttle via this program.

In the early 1980's Goddard will replace most of its present worldwide ground tracking network with two large synchronous communication satellites, the Tracking and Data Relay Satellite System (TDRSS). These two satellites, which will greatly increase spacecraft communications capability, will provide coverage for spacecraft orbiting below 5,000 kilometers. Higher orbit satellites will be serviced by the reduced ground tracking stations. All data passing through the TDRS satellites will be collected at the NASA White Sands, New Mexico, facility where it will be formatted for high speed transmission to users for processing and distribution.

Our interferometry work continues with the development of the Mark-III Wideband Very Long Baseline Interferometry (VLBI) System which has centimeter accuracy. Measurements made in 1977 with our Mark-I system between the Haystack Observatory and Owens Valley, a baseline of 3900 kilometers, indicate about 3 centimeter accuracy. This is about one part in 10 to the 8th. We hope to do much better than this with the new system. The 3 centimeter measurement is the most precise transcontinental length measurement ever made. Transcontinental first order surveys using standard terrestrial surveying techniques are precise to one part in 10 to the 6th. Supporting our VLBI activity is of course our Goddard developed hydrogen masers which offer stabilities of parts in 10 to the 14th.

Oue new hydrogen masers which will be available shortly, will have stabilities of a few parts in 10 to the 15th. These masers will be under microprocessor control for remote monitoring of maser performance and remote control of maser operation such as automatic cavity tuning, synthesizer control, and zeeman frequency measurement. The new masers are transportable like our NP series but are more rugged and have longer battery life. The masers will be available for extensive testing during early 1979.

In support of all this scientific activity during the past twenty years has been the frequency and time community. Much of the scientific data that has been collected could not have been properly interpreted and correlated without precise frequency and time sources. Goddard personnel have, over the years, evaluated and used many techniques of time transfer including HF, VLF, dual VLF, radio navigation systems, portable clocks, television, and the use of satellites.

We are now looking to the future use of our own Tracking and Data Relay Satellites for submicrosecond timing of our Network and other facilities. During the mid 1980's, we also plan to make use of the Global Positioning System (GPS) for submicrosecond timing of our laser ranging network and special projects.

As our scientific programs have become more and more sophisticated, the development of precision frequency and time sources has kept pace. We have seen timing requirements increase from ten milliseconds worldwide during the Vanguard and Mercury flights of the late 1950's and early 60's to tens of microseconds in support of the Apollo program and scientific satellites, to today's microsecond requirements for laser ranging, scientific satellites, Space Shuttle and geodesy measurements. Goddard's timing requirements have increased about an order of magnitude per decade, and we foresee that during the 1980's and 90's the experimenters will need tens of nanoseconds timing which boggles the imagination. Likewise, frequency requirements have advanced from a few parts in 10 to the 8th in the 1950's to the present capability of parts in 10 to the 15th.

Only you people here today can possibly envision what will be required over the next twenty years and on into the 21st century. It is meetings such as this that stimulate the thoughts that bring forth the ideas that develop the systems of tomorrow. I encourage each of you to vigorously continue with your efforts. President Carter, in a message to the Nation last month, pointed out his desire that the United States play a prominent role in fostering space cooperation with other countries. It is meetings such as the PTTI that helps us expand our international activity. I welcome and encourage the participation by other nations in the PTTI.

I want to express my appreciation to the other organizations that cosponsor this meeting; Admiral Fowler and Admiral Smith from the Naval Electronic Systems Command; Captain Smith and Dr. Westerhout from the Naval Observatory; and Captain Henifin and Dr. Berman our host from the Naval Research Laboratory. I thank you for this opportunity to speak with you this morning.

Page intentionally left blank

OPENING COMMENTS

RAdm. G. H. Smith
Vice Commander, Naval Electronic Systems Command

RADM. SMITH: Welcome to the Tenth Annual PTTI Meeting. It is a distinct pleasure for the Naval Electronic Systems Command to once again cosponsor this event.

There is a considerable concern within the Navy about the increasing number of systems requiring precise time and time interval. This concern unfortunately is not matched by an understanding of PTTI.

As an example, last spring there was a requirement to send a cesium clock to southern Europe to update the Verdin systems there. The standard we were sending, of course, required power on it more or less continually.

In order to keep power on the clock during the trip, arrangements were made with Pan American Airlines for the clock to travel first class. Unfortunately, one month earlier new guidance from the Secretary of the Navy put severe restrictions on first-class travel. If you weren't traveling in a stretcher or in a wheelchair or some other such thing as that, it was impossible to travel first class; and it certainly was not possible for an inanimate object to travel first class.

In trying to cope with this problem, we finally agreed that since we couldn't have a thing go first class, we had to at least make it appear that it was a person. So several messages were exchanged having to do with three people traveling first-class to Athens, Rome, and Naples. One of those persons was Mr. C. Clock. It seems that the only person that can give approval for this is the Secretary of Defense, personally. We spent more than two months trying to get approval.

Finally, I took a call from someone in OSD who said, "I don't understand what this is all about, anyway. Why don't you just listen to that Washington radio station and get one of those time checks?"

It is not surprising that you are not very well understood, because you are in the time business, and I challenge you to define time. You probably have argued about this among yourselves, time and time again; but about 40 years ago Albert Einstein defined time. Of course, time was very important to him, in those days. He defined time as "a succession of nows." But it is all relative, anyway.

For the past three years, we in the Navy have been trying to cope with our requirement for the future. We have finally come to grips with it; and it is very simple.

What we need, downstream, is a worldwide dissemination system, and a system on our platforms to distribute time throughout the ship. We don't need to go to picosecond timing in a foreseeable future. We need a way to cost-effectively get the time around. That is what we need.

It is true that our time requirements are becoming more stringent. But for the foreseeable future, they center on those time requirements which are governed by the Transit Improvement Program for navigation, and the GPS Navigational Satellite System; in the wildest stretch of the imagination, that is perhaps 10 nanoseconds.

We must also stop the proliferation of different types of clocks simply because we can't afford it. One solution might be to get a single, very reliable, type of clock that everybody can use. We haven't been able to do that.

We also frequently have a major problem setting our requirements. The first thing an engineer who has a system does, when he sets the requirements, is to look for what is available; when he gets the best he can find, that suddenly becomes the requirement.

Well, in the interest of time, I will make this short. You have three days of detailed technical sessions ahead of you. Please get the most out of the meeting and good luck to you. Thank you.

OPENING COMMENTS

Captain Joseph C. Smith

Superintendent, U. S. Naval Observatory

CAPT SMITH: Dr. Westerhout, Ladies, and Gentlemen: A little over one year ago I spoke at a symposium commemorating the discovery of the satellites of Mars by Asaph Hall of the Naval Observatory. During that talk I stated that I was convinced that there were still things to learn about our solar system, that it was evident that we did not know everything concerning the satellites of the planets nor did we necessarily know about the existence of all the satellites. Less than one year later, I had the honor to announce the discovery of a satellite of Pluto by one of our staff, Mr. James Christy.

While I do not claim yet to be a prophet I believe we can equally make predictions regarding the new and the unknown in the field of timing. Discoveries await which will be equally as exciting as what we have done in visual astronomy.

We are embarked now on a radio astrometric program utilizing the radio interferometer at Green Bank, West Virginia in collaboration with NRL, which promises not only a very precise means to determine UT and Polar Motion, but will also, I am convinced, shed new light on Quasar positions, Quasar physics, and a host of other areas. Our timing experiments in several areas with other individuals and organizations also offer promise for future development in timing as applied to navigation and communications.

Our efforts in time transfer are centered on two developments: A timing receiver for the GPS system, which will eventually allow much better timing world-wide; and the development and performance evaluation of mini portable clocks to cut costs of portable clock visits allowing more to be made. We have intensified our utilization and study of hydrogen masers as time keepers as well as other ways to meet the present and future requirements of the Navy, DoD, and other users.

Most people assume time as a matter of course, yet I am reminded each day of how important timing is to almost every aspect of our modern day Navy life. Electronic systems, navigation systems, communications systems, weapons systems, radar systems, missile systems, computer systems, and many others are going to be increasingly dependent on time. If our platforms of the future are going to be able to do their job, either singly or collectively, such as found in task force operations, a timing system is of the essence. I would like to reinforce Admiral Smith's comments regarding timing and distribution: It must be affordable, economic in operation and capable of world-wide use.

Thus, the challenge is there; there is much more we should know; there is much more we should do. This field is essentially a new one in regards to unprecedented precision and the scope of applications and involvement. There are many new frontiers waiting to be challenged. I am certain that this time next year we shall all be able to say: progress has been made. The amount of progress made rests with you. We of the Naval Observatory are both pleased and proud to be a part of this effort. With your help I am sure we shall be able to say we have kept pace with the clock.

I would like to put in a commercial at this time: For the correct U.S. Naval Observatory Master Clock Time you may call 254-4950, or AUTOVON 294-4950.

Have a good meeting. It is a pleasure for me to see you again.

INTRODUCTION TO "SYSTEMS APPROACH"

Gernot M. R. Winkler
U.S. Naval Observatory
Washington, D.C.

DEFINITION AND OVERVIEW

Systems and timekeeping have several "interfaces":

First, all systems are subject to processes, or better, processes are the essential aspect of systems. In all these processes, time is the one general, abstract parameter which relates the state of these processes to the state of all other systems. Time is, therefore, the universal system parameter (1).

Second, since time is the universal system interface, the use of clocks in systems is increasing, commensurate with our increasing demands for precision in systems interfacing. We now distinguish a class of "time ordered systems" (even though in principle all systems are time ordered) in order to emphasize the precision aspects which necessitate the use of precision clocks (2).

Third, all clocks are of necessity systems; systems in which we attempt to repeat the same processes as identically as possible so that we obtain a uniform time scale. Therefore, any system could be used as a clock, albeit, not a very good one, e.g. we ourselves are systems, i.e. clocks, and one can read the time off our faces!

For these reasons, since clocks are systems, and are part of systems, a general overlook and introduction to the systems approach has been suggested. This appears the more appropriate since some of it, the most critical aspects of the subject, fall somewhat outside the purely engineering frame of mind; the subject is indeed trans-disciplinary. One may even be tempted to claim that these most general, but strategically essential aspects of systems belong to philosophy in its prope sense rather than to any specific technical specialty (3).

Exhibit A attempts to sketch the scope of what is known today under various names such as cybernetics or General Systems Theory (GST), both meaning more or less the same (4).

What is a System?

A system is a group of interacting elements or subsystems which is organized for a purpose. This purpose is clearly external or imposed in the case of artificial or synthetic systems. In natural systems, the purpose is inherent as a heuristic principle (5). The capabilities of a system are ofte clearly reducible to those of its elements; this is the rule for technological systems of up to rathe remarkable sizes and complexities. However, very complex systems (10^8 is a most superficial dividing mark for inter-acting elements) begin to show aspects which are qualitatively new and in principle irreducible to the qualities of the elements (6).

This is partly due to the synergistic interaction of the subsystems and partly due to principle limitations of our intellect. Any intellectual process as we know it is inherently an abstracting process with an incredible and unaccounted amount of data compression at every one of its many stages. Even our sensory perceptions are the eventual results of a most radical selection from the flood of primitive "inputs". In the complex systems environment, however, these ignored, or beter, unknown aspects of the systems elements inevitably come to the fore and play unexpected roles. Then we are awakened to the fact that nature is (at least for us which are its "subsystems

inexhaustible (7). Such ignored parts must appear to us as an irrational part which is irreducible (see Exhibit B #5). This, then, is the origin of "emergent" qualities (8). In our application here, where we consider systems of clocks, the question will, therefore, have to be: What are the emergent qualities of clock systems? We may only hint at this point that one of them is the class of advantages of having a co-ordinated system of clocks with benefits obtained for each user while contributions are made to the timing community as a whole (9).

SYSTEMS, A SUBJECT FOR THE ENGINEERING INTELLECTUAL:

A most important decision at the beginning of any system consideration is to define the level of abstraction at which the system is to be treated. One will hesitate between the "Scylla" of elegant oversimplification and the "Charybdis" of unmanageable complexity and confusion. One may remember, though, that as a typical scientist, one may be wont to miss the forest because of the many interesting trees and also, that the real purpose of any system approach is in fact a view and an assessment of the forest.

But, if we are to lean towards data economy, then we must compensate this paucity with the quality of our concepts. What we wish to stress is that ingenious intuition is really indispensible for the discovery and the judgment of the intellectual tools to be used, i.e. the strategies, goals, concepts and hypotheses. Now ingenuity is something which can't be directed to appear on time. One has to invite it, but the humble waiting for the illuminating insight is worth almost any risk because the truly ingenious concept is literally invaluable (10).

Competent management (in its usual context) is a necessary, but for systems work, not yet sufficient condition. As we hinted before, it is the undocumented, even subconscious experience on which much ultimate systems performance will depend (11). One must therefore suggest that the lack of direct, intimate, personal bench experience of many systems engineers is in the long run very, very expensive.

The two reasons for this are: the need for an environment conducive to create thought and invention, and the need to train judgment and intimate experience. These should make it imperative to start with pilot projects, let people "fiddle around" a little, before serious consequential steps such as specifications writing can be taken.

One can sense at this point that a certain generosity, a large frame of mind, is a desirable systems qualification - but how could one bring anyone to a deliberate expansion of the soul? (Without inflation, of course!)

Now, in going back to the level of abstraction, we have here in these few minutes, no choice but to go for the pinnacle: Exhibit B can of necessity only be a lofty sketch, and I may be forgiven for putting forth such conjectures as verities, almost as if they could be proved by my affirming them on oath. But, these are serious matters and they represent some of the basic things one finds as a result of great complexity.

Some of Exhibit B is purely metaphysical such as #7 (which was already mentioned by Lao-tsu (12)), but some are quite plausibly related to more basic ideas: Item #6 on the list is important because "systemic" measures are disastrous. But, it is also important to understand why that must be so.

They are disastrous because they prevent the system from eventually correcting the real cause of the disturbance. The internal dynamics is being distorted by systemic measures in the very direction which makes the system's state ever more precarious; it is the direction leading into a sudden, catastrophic system collapse (13).

They are most interesting because the study of the effects of systemic measures also gives insights into the connections between GST and "information", correct information and not "noise". We must also include in the importance of information the willingness to use it. This, however, is often prevented by preconceived notions or "policy". The example of economic systems and their inflations is notorious (14). Such a hypothetical system of 10^8 subsystems, each interacting with (let us assume) 10^4 others is a system of great complexity. We may take as a coarse measure the number of possible interactions per unit time, e.g. $\sim \binom{10^8}{10^4}$ (15).

Let us now only sketch a few more points which, while they are obvious, are obviously not being acknowledged easily:

THE EXISTENTIAL LEAP IN THE DARK

Whatever we do, we are forced to act in ignorance of the total consequences of our actions - life is a leap in the dark. In systems design, however, we are not totally ignorant. We have module specifications and user experience with available modules as "shelf items" for building systems. But, the crucial point is this: If we start from overall systems specifications, we will find that many of these shelf items won't do the job. In order to stay within specifications we must custom design, build, test, debug, evaluate, change and produce many of our modules from scratch. (An item is not mature before several hundred have been operated in a system environment and the problems fed back to the debugging process). This, however, is a very expensive enterprise where one's resources may quickly disappear in guerrilla warfare with Murphy's Law (16).

The lesson is that very successful systems must not come into existence via the "grand scheme" with detailed dream-specifications. One rather has to start small (pilot project) and one should define only general system goals (17). But, one can require that all subsystems (modules) be items which have been in production and used sufficiently long to bring out and correct all their hidden problems.

Such an approach does require a greater resourcefulness from your team, but you induce the application of brain power at the critical start of the project rather than for problem solving in the middle (when you hoped to be at the end). This way you don't have to leap in complete darkness.

Another point is more subtle so let us invoke an appropriately exotic example:

THE SIAMESE CAT

Let us consider the procurement of such a "system". Unfortunately, we cannot order it anymore as it used to be done: "One cat (Siamese) each . . . $50.00". Now, we need specifications and they must be complete or exhaustive - but can they be complete?

As we learn from biology, a cat's specifications are embodied in its genes, a total of about 10^{10} bits of information. Prof. Carl Sagan has actually proposed (18) that we should send these 10^{10} bits of information by radio to an alien civilization. This would then enable them to get cats just like ours here on Earth. Now to be charitable, we assume such proposals to be in jest because they would reveal a complete ignorance of rather vital points in systems engineering and specifications! There would simply be no cat on the basis of these 10^{10} bits unless we could also enable that alien civilization to reproduce exactly the totality of our terrestrial environment in which we mass-produce our cats. (In that case, unlikely as it is, they would also have to select a competent and honest contractor, a hard task, because the also imported rules, as part of the environment, may hinder that!)

We, therefore, conclude that we need much more than the 10^{10} bits - but how much?

Just the network of the cat's 10^{10} neurons has up to about 40 interconnections (synapses) per neuron. A description of it would require an amount of information vastly in excess of our 10^{10} bits in the genes. We get a glimpse of the fantastic magnitude of this information problem but we still could not assemble the system unless we have all the information, or could we (19)? The Siamese cat has helped to shed light on the question of how nature builds such extremely complex and well functioning systems on the basis of clearly insufficient information (20).

It turns out that one of the characteristics of the Siamese cat, the grey color of its fur, is caused by a simple gene mutation. This mutated gene cannot, as the normal gene does, produce the pigment melanine at 37°C. Now strangely this failure also causes a second peculiarity which is not obviously related to it: The cross-eyed look and the specific behavior of this cat. The reason for this is that they also have some of their "wires" crossed, i.e. parts of the optical nerves end up in the wrong cerebral hemisphere (21). This produces the need for compensation with some further changes in brain structure during growth. It also induces different behavior, that is why they look a little cross-eyed.

But, how can the failure to synthesize melanine also cause such considerable changes in the growth and structure of the nervous system? The revealing answer is that the absence of melanine in the early stages of cell division creates a different environment for cell specialization which now follows different paths. What the growing organism does is that the chemistry code only modulates the environment, but the actual task of specific design is executed by the generations of dividing and specializing cells, the specific specialization being induced by the local environment which in turn is being changed by the collective effect of the dividing (and specializing) cells.

That is what is called organic growth - the system is self designing. One can learn from nature because a living system such as the one discussed is the result of billions of selections and modifications. If we want to apply principles which correspond to organic growth, then we must also include substantial redundancy, a large reserve capacity (a large margin of performance) and sufficient feedback and internal control at the lowest possible level. Large complex systems have features which in some ways resemble an organism except that we cannot afford to use nature's wasteful techniques of evolution. Instead, we have to use reason and foresight. We are not yet doing too well in that department because we must remember that the public disenchantment with technology has a simple cause. It is not a consequence of too much thinking on our part. It is often intellectual arrogance which prevents completely thorough intellectual preparations. It is also often a tolerance of sloppy, superficial thinking. Therefore, a truly scientific attitude, most helpful in the face of large problems, is not pride and conceit, but the admission of ignorance and eagerness to learn.

CONCLUSION

As a summary, one may look at Exhibit C. Love of truth means also hate for the half-truth and the meaningless. Humility, preached since Socrates, is often eclipsed by the mistaken desire to project an appropriate image. Patience, finally, is rarely appreciated by eager beavers who believe in frequent transplantings as a stimulus to growth. But, this is not so. To build a competent team requires a couple of years. Any technical management which redirects and reorganizes every year, simply does not know what to do. It is, after all, not gimmicks which make a large system a success, but creative, thorough thinking and a very sustained effort necessary to reduce it to practice.

"Patience, reason, and time make possible the impossible." Friedrich Rückert

EXHIBIT A

MAIN ASPECTS OF SYSTEMS THEORY

(GST - CYBERNETICS)

1. Linear Systems: Feedback, optimal control, stability - systems synthesis from known element parameters. Systems architecture (22).

2. Process Statistics: Estimation theory, filtering and prediction - pattern recognition (23).

3. System Measurements and System Identification: How to characterize system performance abstractly. Links to epistemology (24). Physical causes vs. teleological causes (goals).

4. Information Theory: I. Storage (memory), processing - semantics (meaning) - artificial inte ligence - strategies. (Theory of games) (25).

5. Advanced Mathematical Tools: Infinite dimensional state spaces, nonlinear systems, adaptiv systems; systems modelling and simulation; catastrophe theory. Fuzzy sets, recursive an computable sets.

6. Self Organizing Systems: Emerging properties, synergy.

7. Systems, Their Environment and MAN; the Questions of Policy: Man - system interfaces, "Operations Research" (26).

8. System Reliability: Self repairing and redundant systems. System Life cycles and support (27).

EXHIBIT B

SOME GENERAL SYSTEMS PRINCIPLES

1. If anything can go wrong, it will (follows from complexity which assures that eventually the ever present disturbances will test every weakness). Also, known as Murphy's Law.

2. Le Chatelier's Principle: Systems tend to oppose attempts to change them. (In order to function, however poorly, a system must be in internal dynamic equilibrium in its element interactions; therefore, a new input will immediately bring forth compensating forces) (28).

2a. Corollary: If a system should ever work well, then don't touch it!

3. Systems grow (at least due to the after thoughts) (29).

4. A complex and working system can only evolve from a simple and working system (evolutionary growth).

5. Every very complex system contains irrational features; they are necessary for functioning.

6. Symptomatic cures are useless, they only make things worse. The same is true of "systemic" cures (system-wide, sweeping measures, see principle #2). People who don't understand systems love systemic cures!

7. An overdose of the best measures (beliefs, principles, ideas, etc.) is poisonous. (Also, known as the DIALECTICAL (30) Principle in some ideologies - namely in those who overdo this principle also!)

7a. Corollary: You can't be perfect; any attempt to be, or to make a system perfect will have disastrous consequences. (That should not mean that one must stop at the level of mediocrity!)

8. The behavior of complex systems becomes unpredictable as the complexity increases. (See also principle #5)

8a. Corollary: Extremely complex systems are beyond human capacity to evaluate.

9. A system (composed of subsystems) which can operate well with little internal control and data flow will also be highly resistant to failures and external disturbances. Examples: An atomic clock with excellent crystal oscillator; A time ordered system with largely independent clocks; An organization with properly delegated authorities; A society with citizens of common high moral standards.

10. The main cause of system instabilities are the delays in the action of the subsystems on each other. Very large and complex systems, therefore, have no defined equilibrium state (static) and are, therefore, subject to corresponding process statistics (such as flicker noise, random walk, Pareto distribution, etc.) (31). (See also principle #8)

11. Engineering of systems is a hard compromise between opposing desirable features. Systems optimization must include all aspects and is usually successful only if some flexibility of specifications is allowed (33).

12. Any theoretical treatment of a system has an optimum degree of complexity (such as numbers of parameters, etc.). The optimum corresponds to a "best" separation of essential from accidental features (or of deterministic from random effects) and it is usually better to err on the simpler side (32).

EXHIBIT C

THE CARDINAL SYSTEMS VIRTUES

TRUTH	-	Keep The Coefficient Of Fiction Small.
HUMILITY	-	We Know Very Little, Keep Learning.
PATIENCE	-	Good Things Need Time.

NOTES AND REFERENCES

(1) Time as an abstract parameter for the comparison of different processes is discussed and reviewed in

Marton, L. (ed) (1977) "Advances in Electronics and Electron Physics" Chapter 2, p. 34–45, ACADEMIC PRESS, New York ISBN 0-12-014644-4.

(2) "Time Ordered Systems," Time and Frequency, PTTI Systems – all refer in a general way to the enhanced importance of precise time or frequency in these systems. However, one has to be specific in terminology and should distinguish between (a) equal frequency, (b) accurate (a priori, without calibration) frequency, (c) internal synchronization without accounting of differential propagation delay and/or frame ambiguity resolution, (d) internal synchronization with ambiguity resolution and/or delay accounting and (e) systems which are synchronized or coordinated (i.e. whether the time constant is short or very long) with UTC. The step from (c) to (d) is expensive, from (d) to (e) is not, yet it brings the typical full synergistic benefits of being a member of a user community. An example for (c) is the TV, while network synchronization needs usually (d). See the following for a valid recommendation to take, in this case, the easy additional step (e).

Stover, H. A. (1973) A Time Reference Distribution Concept for a Time Division Communications Network. Proceedings 5th PTTI Conference p. 505–523.

(3) The converse seems also to be true as the following may suggest: We only have to substitute systems for the word "being" (Monas, μοναs) and we can immediately understand in our terms the great Renaissance thinker Giordano Bruno when he postulates the universe as a hierarchy of "beings". Now Monas means something which is simple, a unity; a concept which originated from the introspective experiences of one's soul. But, this unity and separateness is deceptive and also makes his view somewhat inconsistent. The soul is due to the synthesis of the sensitivities of the body which are rooted in the (admittedly mysterious) properties of matter and these are universal and eternal (Spinoza: "Sentimus experimurque nos aeternos esse"). Therefore, there is nothing separate about the soul except the limited memory. Conversely, we may see a system as a Monas (black box) because the system concept is purely heuristic, is an intellectual device to divide the world into interesting and manageable parts with the dividing "surfaces" somewhat arbitrary (and the system "transfer function" could be called its character); in other words, the world is a hierarchy of subsystems. Now some of these subsystems have fates which can be almost independent of the rest of the universe for long times. Stable atoms and particles are such (closed) systems and we attempt to use them as clocks. However, completely closed systems can't be predicted from the outside and we can see here clearly the connections to "causality," "indeterminism" (or freedom) and similar general questions.

(4) The literature on GST as well as on the areas of Exhibit A is now enormous. A few general items follow:

Bertalanffy, Ludwig von (1968) "General Systems Theory." Braziller, N.Y. Also published by Penguin University Books, 1973.

Bertalanffy, L. and Rappaport, A. (ed) (1962) "General Systems" Ann Arbor General Systems Research.

Fox, V. (ed) (1965) "System Theory," Interscience Publication, Polytechnic Press, Brooklyn, N.Y. (Lib CCC#65-28522).

Klir, G. V. (ed) (1972) "Trends in GST," N.Y. (Lib CCC#71-178143) John Wiley.

Gall, John (1975) "Systemantics," Quadrangle/The New York Times Book Company (an amusing, but a little to-be-taken-seriously farce).

Guillemin, Ernst (1963) "Theory of Linear Physical Systems," J. Wiley.

Forrester, Jay W. (1961) "Industrial Dynamics" and

Forrester, Jay W. (1971) "World Dynamics," Cambridge, Mass. (Lib CCC#70-157752). (These last two deal with systems dynamics and model building at a large scale.)

(5) By that we mean the teleological explanations. But teleology is purely a heuristic device in science; in philosophy it may be more, much more, because such teleology can only be grounded in some metaphysics.

(6) There is no simple measure for complexity or size of the systems because such measures would have to reflect not only the number of possible interactions per unit of time, but also take into account the complexity (or the character) of the interacting subsystems. Obviously an organization of 100 people is vastly more complex than a computer with 10^8 transistors - but only if they interact.

(7) For a discussion of "inexhaustibility" under the quantum - mechanical point of view see the concept of "qualitative infinity of nature" in

Bohm, D. (1957) "Causality and Chance in Modern Physics," Van Nostrand.

In our context, however, we offer simply the conjecture: There are only subsystems, no real elements.

(8) Emergence is an important concept in British newer metaphysics:

Alexander, Samuel (1927) "Space, Time and Deity." Dover, N.Y.

Here, however, it is based on the informational aspects of epistemology and is much closer to p. 61-81 of

Broad, C. D. (1925) "The MIND and its Place in Nature." Kegan Paul, Trench, Trubner and Company, Ltd. London.

(9) This aspect is discussed in reference (1) p. 79. But, the paper by S. Stein and F. Walls which followed this presentation gives yet another aspect of an emergent quality, i.e. better performance obtainable from a combination of two different oscillators. This is a most important subject for this conference.

(10) Here again it is hard to overemphasize to an audience so rooted in the real and material world as one is today, the fact that some insight gained by concentrated thought is the basis of everything we do today in R & D. And, conversely, everything will become obsolete through another idea. Even more, one could support yet another conjecture: The larger the task, the more sensitive it is to the effects of creative thought.

(11) This is even true for our laws, for anything documented. Polanyi has pointed out that any system of rules can only be transmitted by tradition, example and training. This is clearly a

precarious process and the information so transmitted is wide open to profound changes albeit from one generation of practitioners to the next. Now it is the bureaucratic ideal to leave nothing to discretion, but to embody everything in rules. But, any rule must be interpreted in a concrete case and this would require another rule. However, let us assume that all existing rules were united into a single code; then this code, obviously, could not contain prescriptions for its own interpretation, i.e., our ideal has a principal flaw. We do eventually depend upon responsible, creative interpretation. See

Polanyi, Michael (1946) "Science, Faith and Society," p. 58. University of Chicago Press. (ISBN O-226-67290-5)

Now if the above is true for a body of laws and regulations, then it is a fortiori true for technology where so much depends on the skills, the tricks and trained judgment of the practitioner. In fact, a trade, even a technology, gets irretrievably lost if not practiced. That means that the great majority of know-how developed every year in our national R & D efforts is completely wasted because of the haphazard way in which support of such groups becomes allocated. It would be wise to implement more long range funding plans of R & D teams by mission oriented agencies, such as the DoD, under this aspect.

(12) Lao-tsu (~500 BC) "Tao-te-ching," Vintage Books 1972, ISBN O-394-71833-X.

This is an antidote to the instinctive managerial chrestomania. Lao-tsu was the first to teach that one sees more by not focusing on a single point; that one accomplishes more by not running blindly after a single goal. Today these basic ideas of Taoism find their corroboration in the organic view and in the concept of organism.

We can consider principle #7 of Exhibit B to be of principal importance for any intellectual treatment, of large technical systems as well as of our social situation at large in which, after all, all our technical systems are embedded. Since our social structure as well as our personal attitude rests on a fundamentally inconsistent system of tacitly assumed values and principles, we try to cope by falling from one extreme application of a certain A to an equally extreme one of another "good" B, after the results of the previous excess become painfully obvious. It is for this reason that successful, i.e. long term beneficial systems engineering can be done well by people who have gained that inner directed, self-disciplined attitude which helps them to abstain (out of wisdom, not interest in near term) from overdoing the "best," i.e. from maximizing the "profits." Hence, the claim about the necessary expansion of the soul; and hence, the problems created for science and engineering by the weakening of the liberal arts in education in the course of the post-Sputnik hysteria. One can safely claim that a clear awareness of the need for balance is the prerequisite for any use of intellectual tools. After all, those are more effective and potentially more disastrous than atomic bombs (which are only their by-product).

(13) That is a subject of great recent interest and has become known as "Catastrophe Theory." See e.g.

Zeeman, E. C. (1976) "Catastrophe Theory" in Scientific American, p. 65 (April).

(14) Another notorious and often tragic example is the complex of unintended effects if a specific management problem such as incompetence in some places is being attacked with general reorganizations. Such measures are of course sometimes needed such as when the organizational input-output is to change. But, as symptomatic cures they are disastrous because as systemic cures they disrupt internal communications. See particularly

Lawrence, Paul R. (1955) "How to deal with Resistance to Change" in Harvard Business Review: "On Management," Ch. 22, Harper & Row, ISBNO-06-011769-9 and

Drucker, Peter F. (1973) "Management," Ch. 48, Harper & Row, ISBN O-06-011092-9.

Organizations are also systems, but even more important, any system effort is of necessity a team effort and the first question always to be asked is: How can we get a competent team organized and motivated? Money is only a necessary, but not sufficient resource.

(15) A large scale system has been defined as a system "whose large dimensionality makes the application of standard analysis techniques infeasible due to excessive computational requirements." A discussion on stability and optimization of such systems can be found in

Mageiron, E. F. (1976) "Topics in the Study of Inter-Connected Systems." Technical Report 664, Div. Eng. and Appl. Phys. Harvard.

On the other hand, a system has also been defined as large "when decentralized control can provide acceptable performance." See

Suri, Rajan (1978) "Resource Management in Large Systems." Technical Report #671, Div. Appl. Sc. Harvard University, p. 185.

We propose to call a system large when it can only be dealt with by statistical measures; and very large when entirely new, irreducible features emerge. By this definition, the U.S. economy is a large, but not yet a very large system (since it can go broke through foolishness, just like a family can).

(16) Once a system is in operation, it is too late to improve poor reliability. The larger the system, the more critical will be error free operation. Now transient errors in computers, e.g., are several orders of magnitude more frequent than actual hard failures. The best way to check for transients is hardware redundancy built into critical points. In our application of time keeping, the soft failures are large clock rate changes and here the only way to filter them is clock redundancy. In general, hardware redundancy gives the capability of continuous systems diagnosis, i.e. the decoupling of reliability from other problems. See also

Percival, D. B. et al., (1975) "Time Keeping and the Reliability Problem." Proc. Ann. Frequency Control Symposium, Atlantic City, N.J. 29/412-416.

(17) The "start small" recommendation has as an alternative the advice to proceed if possible only with dividable efforts. Such efforts are those where any intermediate stages are fully beneficial. Building a system of new, high capacity trunk lines is a dividable effort. If money runs out, the effort expended was not wasted. The GPS program (certainly for time dissemination) is also largely in this class. Benefits do not have to wait until everything is perfect. In contrast, a tunnel is clearly not a dividable effort; neither is a bridge.

(18) Sagan, Carl (1973) "Communication with Extraterrestrial Intelligence." MIT Press, Cambridge, Mass.

(19) If we accept the thesis of inexhaustibility of nature (7), then it is clear that no information will suffice to reproduce a natural system exactly. The case is entirely different for artificial systems because they represent the result of abstract functions as conceived by an intellectual process of finite steps. In this case, the ignored aspects of the material implementation play only a tolerably disturbing role - until we exceed a certain complexity.

(20) While no finite amount of discreet (abstract bits) information may suffice to duplicate exactly our cat (which would have to include its acquired behavior), we can quantify the information which is sufficient to launch the germ cell on its development, i.e. the modification and acquisition of parts of its environment. The information is quantifiable and "simple" (10^{10}) because it is relative to the material in the cell and its environment. This is where we have still hidden the inexhaustibility.

(21) Guillery, R. W. (1974) "Visual Pathways in Albinos," p. 44–54, Scientific American, May (See also lit. cited at p. 144 and Nature 252, 195–199, 1974).

(22) The practicality of digital control loops has opened up a new horizon even for simple straightforward systems such as a quartz clock which is locked to an intermittently available reference. One is not limited anymore to the type 1 vs. type 2 (or 3?) loop question, but can now consider much more sophisticated control, i.e., along the lines discussed in reference (23d). As general background refer to chapter 17 of (27). Of particular interest to time and frequency users is also

Gardner, Floyd M. (1966) "Phaselock Techniques." John Wiley LCCC#66–22837.

An excellent introduction to the advanced mathematics of extremal problems with constraints is

Gumowski, I. and Mira, C. (1968) "Optimization in Control Theory and Practice." Cambridge University Press ISBN 521–05158–4.

(23) a. Koopmans, L. H. (1974) "The Spectral Analysis of Time Series," Academic Press, N.Y. ISBN O–12–419250–5.

b. Box, G. E. P. and Jenkins, G. M. (1970) "Time Series Analysis, Forecasting and Control," Holden–Day, San Francisco, LCCC#77–79534.

c. Morrison, N. (1969) "Introduction to Sequential Smoothing and Prediction," McGraw-Hill, N.Y. LCCC#69–17187.

d. Percival, D. (1978) "The U.S. Naval Observatory Clock Time Scales," Trans. IEEE–IM Dec.

The first two treat random variables after systematic parts have been removed. Morrison deals with estimation of deterministic functions in the presence of additive random errors. Percival discusses the application of ARIMA models (the Box–Jenkins approach) to the treatment of clock sets.

(24) For the treatment of very complex systems, it is necessary to be quite clear about what we mean by "cause," "model," "law," "probabilistic explanation," etc. These are concerns of practical relevance, e.g. why do we have confidence in extrapolations with deductive nomological explanations – causal laws, when we should be very suspicious of extrapolations with a purely mathematical model? For the pure empiricist, the difference is harder to justify than for the rationalist (or for a medieval realist, or Platonist). It is no accident that the treatment of the most basic systems in science, i.e. atomic and particle physics and of the most complex, possibly inexhaustibly complex, systems, i.e., in biology, has brought about an unprecedented concern of those scientists with epistemology. As a very first starter, the following is suggested:

Reynolds, P. D. (1971) "A Primer in Theory Construction," Bobbs - Merrill, Indianapolis ISBN O-672-61196-1.

(25) We can recommend a reference which is of interest beyond the scope of its title because it gives very instructive examples for difficulties of quantification and discussions of "utility," "rational behavior," etc. It is a classic:

Neumann, John von and Morgenstern, Oskar (1944) "Theory of Games and Economic Behavior," John Wiley and Sons, Inc., N.Y. ISBN 0471911852.

(26) Wagner, Harvey M. (1975) "Principles of Operations Research" Second Edition, Prentice Hall, Inc. ISBN O-13-709592-9 is a textbook with many "mind expanding" exercises! To work through these 1000 pages (this author has not!) should give a firm grounding in OR.

(27) An excellent general overview is given in

Giacoletto, L. J. (ed) (1977) "Electronic Designers Handbook," Section 28 (by Trent), McGraw-Hill, ISBN O-07-023149-4. (See also note (15))

(28) A logical corollary to principle #2 of Exhibit B would be: Indirect approaches often work better (because they avoid the system reaction, at least it is delayed into ineffectiveness). The best treatment, of course, would be a thorough understanding of the causality of the system dynamics, but this will be available only in systems of modest complexity. (See also principles #5 and #6 of Exhibit B.)

(29) Since these after-thoughts seem to be inevitable, it will be wise to make provisions for them, i.e. to allow for expansion in the basic architecture.

(30) Unfortunately, Hegel, the inventor of the modern sense of "dialectic," confused an epistemological with an ontological issue. In explaining things, the isolation of opposing ideas is an important intellectual device. In the ensuing discussions, the extremes find a synthesis which is a better explanation than either of its components, the thesis and antithesis. But, this purely explanatory aspect has nothing to do with a common S property - the tendency to invert things. Examples are in the bible (1 Cor. 1,27); Plato (Laws 4, 705); Aristotle calls it "Reversal of roles" ("peripeteia" in Poetica VI, 18); and Toynbee in his "A Study of History" (Oxford University Press 1972, ISBN O-517-179415) in Chs. 22 and 23 gives historical examples where failure was caused by early success.

In modern terms, Russell (Authority and the Individual, Simon and Schuster, N.Y. 1949, p. 9 and 66) suggests that Security is a misguided goal in social perfection. In the most abstract terms, one can state it as follows: An unbalanced, single minded application of an A will be inverted through the S process into its eventual opposite, with the effect as if A^{-1} would have been applied.

Examples: A large system designed with sole emphasis on cost will, through breakdowns and deficiencies become more expensive than a reasonably generous design. A timing system designed with maximum emphasis on performance margins will, through the use of "racing horse" type clocks, become very touchy, personnel dependent and generally unreliable.

(A bureaucratic system designed for centralized efficiency with tight control and detailed procedures will become unresponsive to needs, extremely wasteful and an easy target for the most blatant corruption (see also (11).)

Now the tragic confusion, mentioned above, was inherited by Hegel's materialist successors who always see in such instances "contradictions" instead of real system reactions. Contradictions are a purely logical affair - we create them if we are confused and inconsistent. And, a synthesis may come in discussions, but the system will not find it - they will oscillate wildly. Contradictions are only in our theories, never in the phenomena. Those will go their way, whether we see "contradictions" or not.

(31) Mandelbrot, B. (1963) "New Methods in Statistical Economics," J. of Political Economy 71, 421 makes the point that large systems often exhibit statistics with some internal correlations ("Flicker noise" in clocks, "random walk" and "Pareto" distributions in economic systems). Such systems are hard to investigate because in the presence of such behavior, a basic deterministic performance cannot be discerned, i.e. perceived structures may be due to chance. We claim that this must be expected in highly complex systems with delays (which may be integration effects such as, e.g. credit expansion takes a long time until its inflationary effects become noticeable).

(32) This is variously known as Occam's Razor (do not use more concepts than necessary); Mach's principle of economy of thought (the simplest theory is the best); the principle of parsimony; etc. In primitive terms, it means that one should use the simplest mathematical model for smoothing because otherwise one accepts random components as part of the deterministic model which is disastrous for any extrapolation! See also the interesting numerical example of finding the degree of a noise - contaminated unknown polynomial in

Scheid, F. (1968) "Numerical Analysis," Ch. 21, p. 250. Schaum's Outline Series, McGraw-Hill, N.Y.

(33) System Optimization (principle #11) has also aspects which must be considered under the "Reversal" principle (#7). The "Tragedy of the Commons" syndrome (see Hardin in Science 162/p.1243, 1968) can be generalized in so far as optimization of the subsystems without regard to the system leads to poor performance if things work at all. This is so because "the whole is more than the sum of its parts".

Page intentionally left blank

COMPOSITE OSCILLATOR SYSTEMS FOR MEETING USER NEEDS FOR TIME AND FREQUENCY

S. R. Stein and F. L. Walls
Time & Frequency Division
National Bureau of Standards
325 Broadway
Boulder, CO 80303
(303) 499-1000, ext. 3224, 3207

ABSTRACT

Frequency standards are used in most navigation and telecommunications systems to provide a long term memory of either frequency, phase, or time epoch. From a systems point of view, the performance aspects of the frequency standard can be weighed against other systems characteristics, such as overall performance, cost, size, and accessibility; a number of examples are very briefly reviewed. The theory of phase lock and frequency lock systems is outlined in sufficient detail that one can easily predict total oscillator system performance from measurements on the individual components. As an example, details of the performance of a high spectral purity oscillator phase locked to a long term stable oscillator are given. Results for several systems, including the best system stability that can be obtained from present commercially available 5-MHz sources, are shown.

INTRODUCTION

Frequency standards are used in most navigation and communications systems to provide a long term memory of either frequency, phase, or time epoch. These quantities form a hierarchy, so that, in general, systems which depend on minimum time dispersion for their operation place more difficult requirements on the system oscillator than those which only require phase coherence or frequency stability. We will begin by very briefly examining some of the major applications of frequency standards and consider some of the possible tradeoffs between oscillator performance on the one hand, and the performance of other system components on the other.

Doppler Radar:

The simplest radars are CW Doppler devices which determine the radial velocity of a target by homodyne detection of the return signal with the

transmitted signal. In normal operation, the range and detectibility of targets using such devices is limited by the phase noise of the oscillator. Phase noise on the carrier* in the Doppler band of interest is reflected from nearby "clutter", having very large cross section, and cannot be distinguished from the Doppler shifted carrier reflected from a small distant target: Decreased phase noise therefore results in extended range and/or better contrast, without resorting to more complex signal analysis such as time delayed cross correlation. Very low phase noise can be obtained with a high drive level quartz crystal oscillator. Devices with white phase noise below 10^{-18} rad^2/Hz at a carrier frequency of 100 MHz are commercially available. However, such oscillators have much degraded long term performance compared to low drive level quartz oscillators or atomic frequency standards. Consequently, in order to satisfy spectral purity requirements and long term stability specifications it may be necessary to adopt a systems approach to the oscillator, e.g., a high drive level crystal oscillator locked to a low drive level crystal oscillator or an atomic clock.

Doppler Navigation:

Present day navigation systems both for satellites and for deep space tracking utilize two-way or coherent measurements. A signal is transmitted to the spacecraft where it is transponded for Doppler detection on Earth. The velocity error is proportional to the frequency change of the oscillator during the round trip to the spacecraft and establishes a requirement on the stability of the frequency standard.

Coherent Doppler tracking has the disadvantage that a significant amount--approximately one-third--of the time is spent tracking the spacecraft and this requires the largest, most accurate radio antennas available. The inclusion of an onboard frequency standard eliminates the need for the large antennas. By differencing two one-way Doppler signals, it would be possible to use smaller antennas for shorter periods, thus considerably decreasing the initial capital expenses. The stability requirements of the spacecraft beacon are quite modest; the range information is contained in the differential Doppler signal, causing the noise of the onboard oscillator to cancel to first order. However, to achieve tracking accuracies which are desired for the 1980's--less than 10 cm range error and less than .05 rad angular error--it will be necessary to synchronize independent clocks at the ground stations to better than 1 ns. Since the best existing commercial clocks cannot achieve 1 ns time dispersion for more than one day, either daily resynchronizations or new clock systems will be required.

*See Appendix A for definitions, specifications and general discussion of phase noise.

Geodesy:

The determination of baseline coordinates over geodynamically interesting distances is being done using remote very long baseline interferometry (VLBI) stations. Quasars are used to establish a sparse grid system which may then be filled in by satellite radiointerferometry. The role of frequency standards in VLBI is to establish phase references at each station which are coherent with each other for the duration of the measurement. The standards are used to independently determine the phase of the received signal, thus permitting subsequent cross correlation of the signals. The maximum duration of the data stream which may be cross correlated is small--approximately one-tenth--compared to the correlation time of either reference oscillator. The correlation time, τ_c, of an oscillator is defined so that the integrated phase noise for frequencies greater than $1/\tau_c$ is one rad^2, that is (Appendix A)

$$<\phi^2>_{1/\tau_c,\ \infty} = 1$$

For an oscillator at frequency ν_o whose long term fractional frequency stability,** $\sigma_y(\tau)$, is dominated by a flicker noise level, σ_y(flicker), the coherence time is $1/(2\sigma_y(\text{flicker})\nu_o)$. For example, let σ_y(flicker) $= 10^{-14}$ and $\nu_o = 10^{10}$ Hz; the coherence time is 5000 s and the maximum duration of the data stream is 500 s. The primary tradeoff is time spent for resynchronization versus a more elaborate clock system.

When satellite signals are used for geodetic baseline measurements, there is a tradeoff between antenna and oscillator performance. If a suitable phased array antenna can be built which hops from one satellite to the next once each second, then frequency standards at the independent ground stations having stability of only 5×10^{-11} are required for sub-decimeter accuracy. However, if dishes must be used, the cycle rate between satellites would be approximately one hundred times slower and the oscillator would need to be one hundred times more stable.

Time Code Navigation:

The Global Positioning System will function by the transmission of time and position data from an ensemble of satellites. By observing four satellites, the observer can solve for his three position coordinates and the time. Since the solution depends on the range to a satellite being proportional to the time of flight of the signal, the errors in the coordinate solution are directly proportional to the time dispersion of the onboard clock. The satellite clocks must be regularly resynchro-

**See Appendix B for a precise definition and discussion of $\sigma_y(\tau)$.

nized before the range error due to time dispersion exceeds the system accuracy specification.

Fig. 1 shows the time dispersion of a clock, using optimum prediction

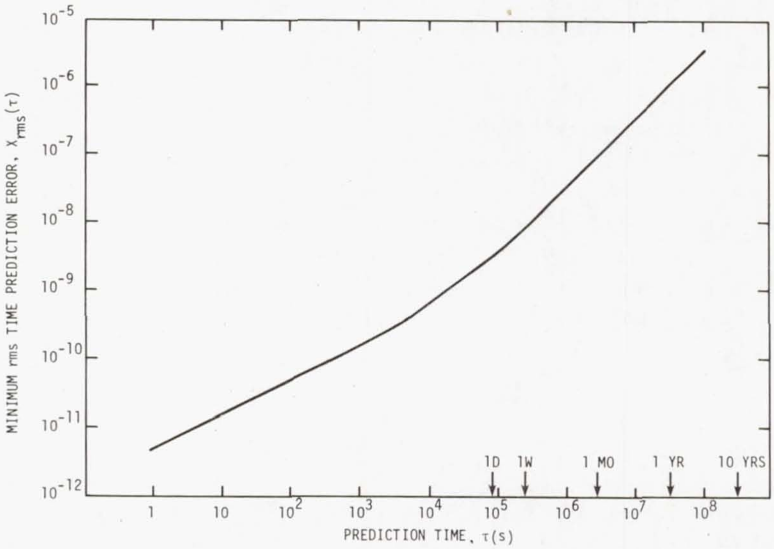


Fig. 1. Time dispersion of a clock with stability $\sigma_y{}^2(\tau)=(5\mathrm{x}10^{-12}\tau^{-1/2})^2 + (3\mathrm{x}10^{-14})^2$, using optimum prediction techniques.

techniques, due to white frequency noise and flicker frequency noise;[1] the assumed time domain stability is $\sigma_y{}^2(\tau) = (5 \times 10^{-12}\tau^{-1/2})^2 + (3 \times 10^{-14})^2$. The rms prediction error, $x_{rms}(\tau)$, is $\phi^2(\tau)^{1/2}/2\pi\nu_o$ where $\phi(\tau)$ is the difference between the oscillator's actual phase (time) and the predicted value after a time interval τ. For the assumed oscillator performance

$$x_{rms}(\tau) \cong (5 \times 10^{-12}\ \tau^{1/2})^2 + (\frac{3 \times 10^{-14}}{\ell n\ 2}\ \tau)^2$$

and the value of the rms prediction error at one day is 4 ns, almost entirely due to the flicker frequency noise. After a sufficiently long time this model will no longer apply because deterministic effects will dominate over the random noise terms. The dominant deterministic term results from a lack of knowledge of the true frequency drift and causes time dispersion which is quadratic in the prediction interval, i.e., $x_{rms}(\tau) = (D/2)\tau^2$ where D is the frequency drift per unit time.

The inclusion of a frequency standard on board a satellite or space vehicle results in several favorable features. In the case of GPS, it permits an unlimited number of users to simultaneously, passively access the system and eliminates the need for the user to have a transmitter

and to track the satellite. However, to satisfy these requirements in a system which is capable of 3 m (10 ns) accuracy 10 days after the last satellite clock resynchronization requires a spaceworthy clock with performance which has only been demonstrated in the laboratory.

Communications Systems:

There is a close relationship between the performance of oscillators and the communications systems which utilize them. Phase coherent systems require carrier synchronization, that is, a precise knowledge of the transmitted carrier frequency and phase. In addition, symbol, word and frame synchronization are usually needed. In some situations, phase coherent techniques do not apply. This is necessarily the case, for example, when the transmitted power is limited and the range is very great. The finite width of the carrier prevents unlimited narrow banding of the receiver in order to compensate for loss of signal-to-noise ratio.

The stability of the transmitted signal has similar effects on both incoherent and coherent systems although the latter are more sensitive. Some typical examples are: (1) The time to acquire frequency or phase synchronization is proportional to the maximum frequency error of the transmitter divided by the receiver bandwidth. (2) The probability of bit errors increases as a result of noisy carrier reference or synchronization signals. (3) The transmitted carrier power decreases as a result of phase noise power. (4) The required channel bandwidth and deadbands must accomodate the long term instability of the transmitter.

Need for Composite Oscillator Systems:

One result of an overall system analysis will be the specifications for a frequency standard needed to provide reference signals. If the specifications cannot be met by a commercially available device, then it may still be possible to meet the requirements with a system composed of more than one device.

Oscillator systems are useful because the operating conditions which optimize one performance aspect of a simple oscillator are normally unfavorable to other aspects. For example, in a quartz crystal oscillator, high power dissipation in the resonator is necessary to achieve the best short term stability (averaging times less than .1 s) but, due to the piezoelectric nonlinearities, much lower power levels are required for the best long term stability and lowest drift.[2] However, a system can be constructed consisting of two devices, one optimized for long term stability and the other for short term stability. The system can have a single output which has the best performance of the two devices.[3] A systems approach is also useful for optimizing other aspects of oscillator performance. Systems can be used to provide power gain after frequency mulitiplication, to provide filtering functions which are not

possible with passive devices and to provide unusual combinations of properties such as high tuning rate combined with superior long term stability. Equations which permit one to predict the noise performance of a system comprised of previously measured components are developed and several examples are evaluated.

One system will be evaluated in detail, both theoretically and experimentally. It is an extremely important example because of its wide applicability and the fact that it can be easily constructed from commercially available components. It consists of a quartz crystal oscillator having state-of-the-art spectral purity which is phase locked to a second quartz crystal oscillator having state-of-the-art long term stability. Data are presented which demonstrate the overall system spectrum and time domain stability as a function of the loop parameters.

THEORY OF PHASE LOCK AND FREQUENCY LOCK SYSTEMS

The general problem is to improve the stability of a voltage controlled oscillator (VCO) by locking it to a reference which has better performance over some range of interest. Two types of feedback loops are normally used: a frequency lock loop is required when the reference is a passive resonator: a frequency lock loop (FLL) or a phase lock loop (PLL) may be used when the reference is an active device producing its own output signal.[3] It is shown below that the same equations can be used to determine the stability improvement, independent of the type of loop.

Fig. 2 shows the general features of the feedback loop. The reference frequency is Ω_r and deviations from the nominal are denoted $\Delta\Omega_r(t)$; the

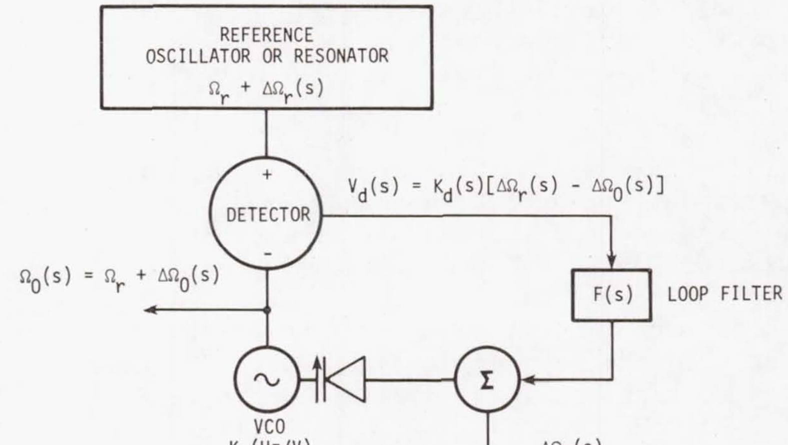


Fig. 2. Phase or frequency lock loop.

VCO has frequency Ω_r and deviations $\Delta\Omega_o(t)$. It is compared with the reference in the device labelled detector whose output is $V_d(s) = K_d(s) [\Delta\Omega_r(s) - \Delta\Omega_o(s)]$, where $\Delta\Omega_r(s)$ and $\Delta\Omega_o(s)$ are the Laplace transforms of the corresponding frequency deviations. The phase and frequency transforms are related very simply by the expression

$$\phi(s) = \Delta\Omega(s)/s$$

which makes it possible to distinguish between the FLL and the PLL entirely through the functional dependence of $K_d(s)$. For the FLL, $K_d(s)$ is a constant, K_v, while for the PLL, $K_d(s) = K_\phi/s$ where K_ϕ is a constant. Thus the detector output for the PLL is proportional to the phase difference between the VCO and the reference, $V_d(s)=K_\phi[\phi_r(s) - \phi_o/s]$. The noise voltage generator $V_n(s)$ represents the internal noise of the VCO. If the tuning rate of the oscillator is K_o(Hz/volt), then the open loop noise of the VCO is given by

$$\Delta\Omega_n(s) = k_o V_n(s)$$

The closed loop performance of the system follows by tracing $\Delta\Omega_o(s)$ around the loop:

$$\Delta\Omega_o(s) = \left[\frac{1}{1+G(s)}\right] \Delta\Omega_n(s) + \left[\frac{G(s)}{1+G(s)}\right] \Delta\Omega_r(s),$$

where $G(s) = K_o K_d(s) F(s)$ is the open loop gain. Assuming that the noise in the VCO and the reference are uncorrelated, the spectral density of the frequency noise obeys the equation

$$s_{y_o}(\omega) = \left[\frac{1}{|1+G(j\omega)|^2}\right] s_{y_n}(\omega) + \frac{|G(j\omega)|^2}{|1+G(j\omega)|^2} s_{y_r}(\omega)$$

where the y's denote the deviations normalized to the carrier frequency, e.g., $y_o = \Delta\Omega_o/\Omega_r$.[4] It follows that $S_\phi = S_y(\Omega_r^2/\omega^2)$, so S_ϕ satisfies the same relation as S_y:

$$s_{\phi_o}(\omega) = \left[\frac{1}{|1+G(j\omega)|^2}\right] s_{\phi_n}(\omega) + \frac{|G(j\omega)|^2}{|1+G(j\omega)|^2} s_{\phi_r}(\omega)$$

$|G(j\omega)|$ generally increases monotonically with decreasing ω, making it possible to draw some general conclusions about the output spectrum of the servoed oscillator. The noise in the reference oscillator and in the control loop is low pass filtered, while the noise in the VCO is high pass filtered. This leads to the most common situation--the output spectrum is dominated by the reference oscillator at low Fourier frequencies and by the VCO at high Fourier frequencies. However, in the event that the spectral purity of the VCO, $S_\phi n(\omega)$, is much worse than the purity of the servo or reference, we see that $S_{\phi_o}(\omega)$ is approximately $S_{\phi_n}(\omega)/|G(j\omega)|^2$. If $|G(j\omega)|$ has a maximum value of 10^4, for example, then the spectral density of the locked VCO can never be better than $S_{\phi_n}(\hat{\omega})/(10{,}000)^2$, even if the servo and reference are much less noisy at that Fourier frequency.

This indicates that $|G(j\omega)|$ should be as large as possible. However, the maximum value is fixed by the rolloff slope and the maximum unity gain frequency which can be tolerated by the short term stability of the reference signal. If the rolloff exceeds 12 dB/octave at the unity gain point, then the loop will oscillate. The required shape of $G(j\omega)$ to reduce the effect of the open loop VCO noise below the level of the reference noise for Fourier frequencies below the unity gain frequency can be determined from the above equation. For example, to transform random walk of phase ($S_\phi \propto \omega^{-2}$) to white phase noise ($S_\phi \propto \omega^0$) requires a single integration (i.e., $|G(j\omega)| \propto 1/\omega$). An analogous result can be derived for deterministic processes which can not be described in terms of spectral densities. For example, the unlocked VCO may exhibit frequency offset $\phi(t) \propto t$, frequency drift $\phi(t) \propto t^2$, or even frequency acceleration. If the parameter to be controlled has an open loop behavior proportional to t^P then the requirement for the closed loop system to have zero dc error is[4]

$$\lim_{\omega \to 0} [\omega^P |G(j\omega)|] = \infty.$$

A critically damped second order loop response is nearly ideal for oscillator system applications. The required open loop gain function is

$$G(j\omega) = \frac{\omega_n^2}{\omega^2}\left(1 + j2\frac{\omega}{\omega_n}\right)$$

in which ω_n is called the natural frequency of the loop.[6]

In the case of the PLL it can be approximated by selecting the filter of Fig. 3a for F(s), while the FLL requires an additional integration in order to have the same overall open loop gain function.

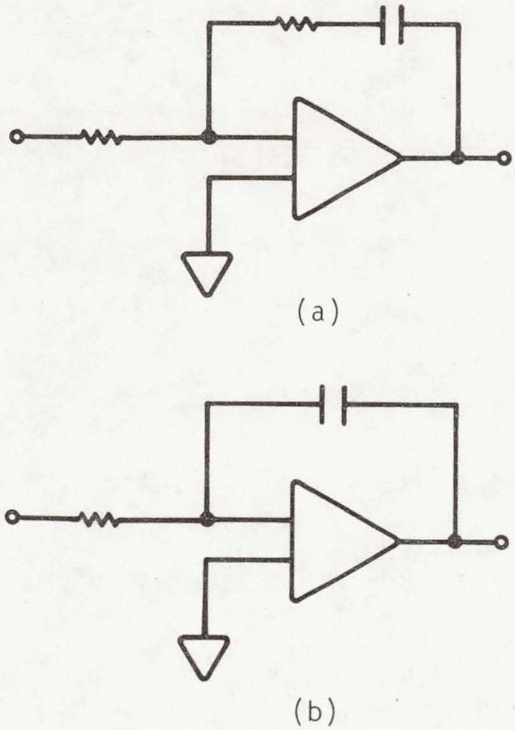


Fig. 3. Filters for phase and frequency lock loops. Filter (a) is used to achieve a second order PLL response while (a) and (b) in series produce a second order FLL response.

For frequencies small compared to $\omega_n/2$, the open loop gain increases at 12 dB/octave. However, because of the breakpoint at $\omega_n/2$, the slope is only 6 dB/octave at the unity gain frequency and the loop is unconditionally stable. Because of the limitations of analog integrators and other circuitry, the PLL, which has an inherent pure integration, is superior in performance where it is applicable. Errors in such a loop produce a phase offset between the two oscillators, whereas errors in a FLL result in a frequency offset. The FLL must be used with passive resonators and can be used to achieve improvments in some PLL characteristics, such as pull in range and acquisition time.

SYSTEMS APPROACH FOR SIMULTANEOUS SPECTRAL PURITY AND LONG TERM STABILITY

There are commercially available oscillators with superior spectral purity and long term stability, but no one device has the best performance for all averaging times. This section will show how to design a

system which has the optimum performance everywhere: Data are presented which confirm the results of the previous section.

The low drift 5-MHz oscillator, used as the reference, is characterized by the measured spectral density,

$$s_{\phi_r}(f) = \frac{10^{-11.3}}{f^3} + 10^{-13.6}$$

while the spectrally pure, 5-MHz oscillator is approximately characterized by

$$s_{\phi_n}(f) = \frac{10^{-10.5}}{f^3} + 10^{-17.1},$$

where $f=\omega/2\pi$ is the Fourier frequency offset from the carrier.

A second order PLL was selected to combine these two devices as a system; Fig. 4 shows the major elements of the circuit. A double balanced mixer serves as the phase detector. It is followed by a 2-pole filter which attenuates the 10 MHz output of the mixer, but has little influence on the loop. The capacitor shunting the operational amplifier and the low pass filter following it limit the noise bandwidth of the loop at a sufficiently high frequency compared to the natural frequency of

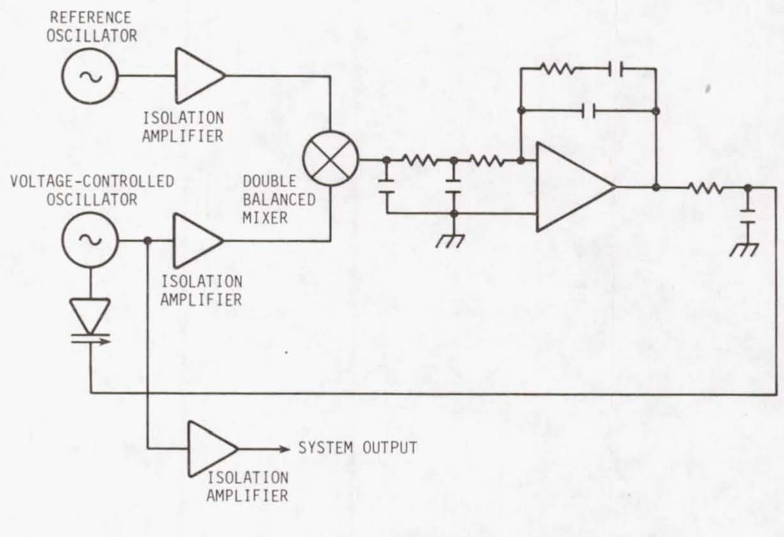


Fig. 4. Second order phase lock loop.

the loop so that they have negligible effect on the loop response. The resistor and capacitor in the feedback path around the operational amplifier are chosen to produce critical damping. The approximate system spectral density is

$$S_{\phi_o}(f) = \frac{10^{-11.3}}{f^3} + 10^{-17.1}.$$

Fig. 5 shows the open loop noise of both oscillators and the closed loop system output for two different unity gain frequencies, 8 Hz (20 ms) and 16 Hz (10 ms). If the unity gain frequency is too high the system spectral purity is degraded by the reference noise; the optimum unity gain frequency is near 8 Hz in this example.

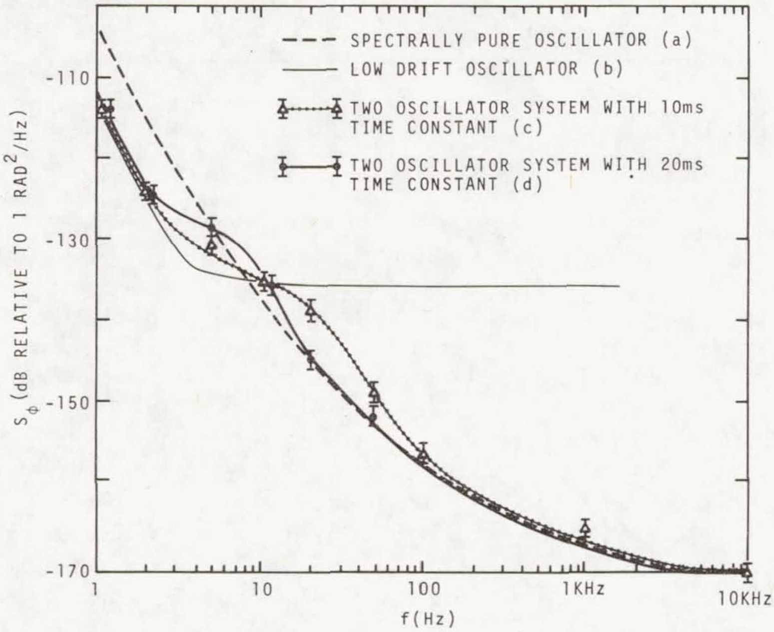


Fig. 5. Spectral density of phase for (a) 5-MHz VCO having excellent spectral purity, (b) 5-MHz VCO having good long term stability, (c) system performance with 16 Hz unity gain frequency, and (d) system performance with 8 Hz unity gain frequency.

The time domain stability is shown in Fig. 6. The triangles are the system stability with 8 Hz unity gain frequency, while the circles are the spectrally pure oscillator open loop stability. The two horizontal lines are the time domain performances calculated from the S_ϕ data of Fig. 5. The line with slope of τ^{+1} corresponds to a drift rate of 3.7 x 10^{-8}/day.

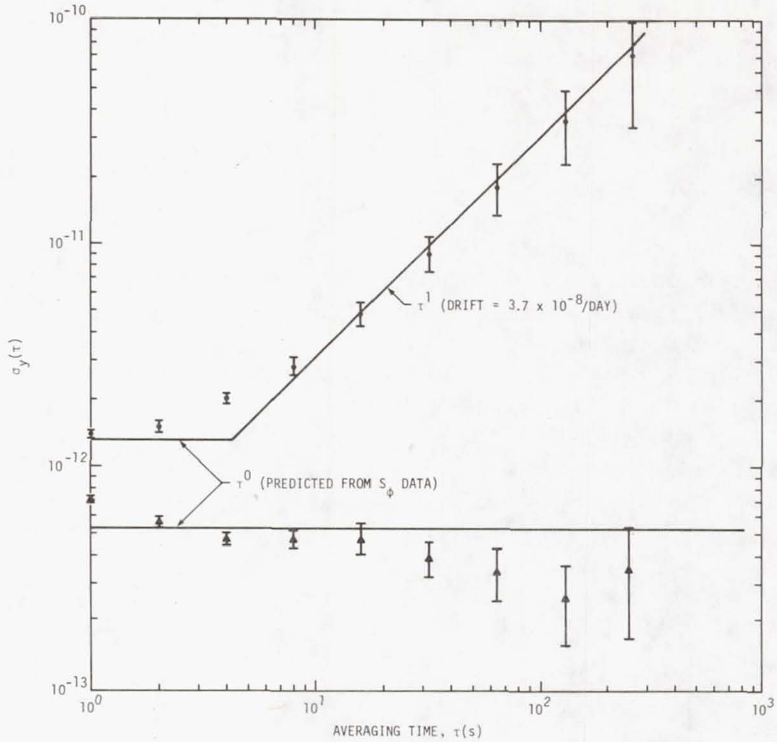


Fig. 6. Time domain stability of the spectrally pure oscillator (circles) and the two-oscillator system with 8 Hz unity gain frequency (triangles).

The results which could be obtained with three other systems are shown in Fig. 7. The stability shown for system 1 results from phase locking three commercial oscillators: A state-of-the-art 5-MHz oscillator provides the short term stability (τ<1s) and is locked to a low drift oscillator for best intermediate stability; the long term performance results from locking this pair to a cesium beam frequency standard. This system has the best performance which can be obtained with commercial devices at 5 MHz.

In the region from approximately .01 s to 1 s, the system stability is better than that of any of the component oscillators. This situation occurs in the case of white phase noise, because the stability of the long term stable oscillator in this region is determined by the very high frequency portion of the spectral density. For this reason, $\sigma_y(\tau)$ curves should be used with great caution when predicting system characteristics.

System 2 is a prototype; a 5-MHz quartz crystal oscillator is frequency locked to a passive quartz crystal. It has the best intermediate term

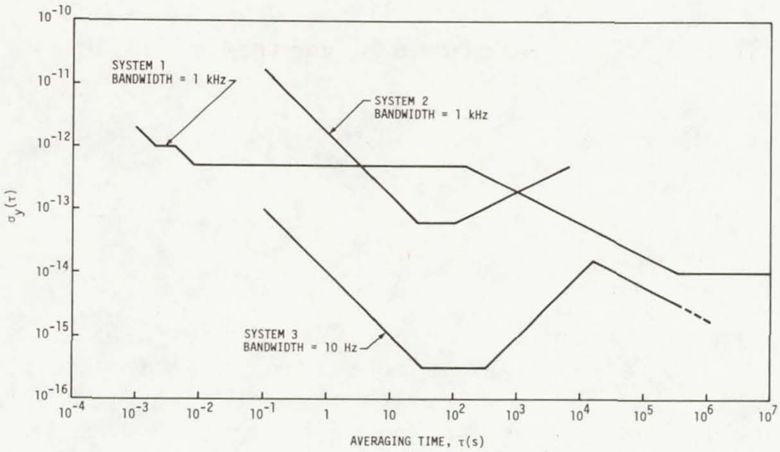


Fig. 7. Time domain stability of (1) a system composed of a spectrally pure quartz oscillator, a low drift quartz oscillator and a cesium clock (predicted); (2) a system composed of a spectrally pure quartz oscillator and a passive quartz crystal (realized); (3) a system composed of a superconducting cavity stabilized oscillator and a passive hydrogen frequency standard (predicted).

stability ever achieved with a quartz crystal, because the passive resonator is operated under conditions which optimize the stability in the 1 to 10^4 s range.[2,3] Eventually system 2 should equal or better the performance of system 1 for all times shorter than 10^4 s.

System 3 shows the results which would be obtained if a superconducting cavity stabilized oscillator were phase locked to a passive hydrogen frequency standard. The system performance is the best which can be achieved with existing devices.[7,8]

SUMMARY

We have illustrated how a systems approach, with its increased degrees of freedom, can provide greatly improved stability performance relative to a single oscillator. In a completely analogous manner an oscillator exhibiting low vibration or radiation sensitivity can be phase locked to a long term stable oscillator to yield improvements in these parameters.

Simple equations which can be used to predict the performance of either frequency or phase lock systems have been discussed. A loop filter for achieving near optimum results was described and experimental results of one phase lock loop system using this filter were presented. Predicted performance curves for some other interesting systems were presented. Using this information, the designer should be able to tailor the performance of an oscillator system to meet the overall frequency stability, accuracy or timing specifications of a navigation, communications or other large electronic system.

APPENDIX A: Phase Noise in Oscillators

The instantaneous output voltage of a high quality signal generator may be written as

$$V(t) = [V_o + \varepsilon(t)] \sin [\Omega_o t + \phi(t)]$$

where V_o and Ω_o are the nominal amplitude and frequency, respectively, while $\varepsilon(t)$ and $\phi(t)$ are random processes representing amplitude noise and phase noise. The objective is to characterize $\phi(t)$. Traditionally the measurements have been described as being performed in either the frequency domain or the time domain. The recommended definition for the frequency stability measure in the Fourier frequency domain is the one sided spectral density on a per Hertz basis, $S_\phi(f)$, of the random process $\phi(t)$.[9] In terms of the spectral density, the mean square phase fluctuations within the frequency band $f_1<f_2$ is

$$\langle \phi^2(t) \rangle_{f_1,\, f_2} = \int_{f_1}^{f_2} df \quad s_\phi(f)$$

Phase noise in oscillators is often expressed as a ratio of single sideband phase noise, per root Hertz to carrier power--$\mathcal{L}(f)$--as a function of Fourier frequency offset from the carrier. $\mathcal{L}(f)$ is related to the spectral density of phase noise of the oscillator by the equation

$$\mathcal{L}(f) = \frac{1}{2}\, s_\phi(f) \quad \text{for} \quad \int_f^\infty df \; s_\phi(f) << 1$$

Figure 8 shows the typical appearance of the spectral density of phase noise of an oscillator. High quality oscillators often exhibit power law dependence of the spectral density. In region I, S_ϕ is typically proportional to $1/f^3$ and the oscillator is said to have a flicker frequency noise behavior; it is probably the result of changes in the values of the frequency determining elements. In region II, S_ϕ is proportional to $1/f^2$ and the oscillator is said to have white frequency noise; it is often the result of thermal noise in the gain element of the receiver. Normally, the amplitude noise is much less than the phase noise in both region I and II. In region III, S_ϕ is constant and the oscillator is said to have white phase noise. This is usually the result of additive thermal noise in an amplifier or some other device. In this region, the amplitude noise is generally equal in magnitude to the

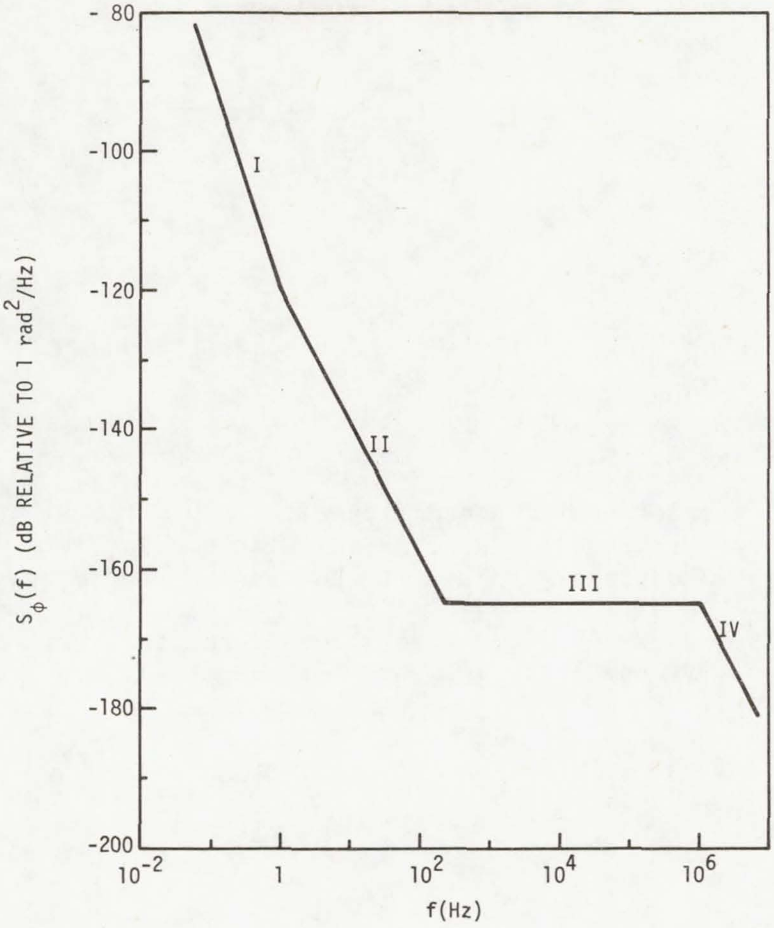


Fig. 8. Typical spectral density of phase for a high quality oscillator.

phase noise. Region IV is usually due to finite bandwidth of the output amplifiers.

APPENDIX B: Frequency Stabiity of an Oscillator

The instantaneous fractional frequency deviation from nominal is defined as

$$y(t) = \frac{1}{2\pi\nu_o} \frac{d\phi(t)}{dt}$$

The recommended definition for the frequency stability measure in the time domain is the two sample, zero-deadtime variance, commonly called

the Allan variance,

$$\sigma_y^2\,(\tau) = \left\langle \frac{(\bar{y}_{k+1} - \bar{y}_k)^2}{2} \right\rangle$$

where

$$\bar{y}_k = \frac{1}{\tau} \int_{t_k}^{t_k+\tau} y(x)dx$$

is the average fractional frequency over the k^{th} interval of length τ, and the angular brackets indicate an infinite time average.[9] It is, in general, also necessary to specify the measurement bandwidth, f_h.

Fig. 9 illustrates the typical appearance of the two sample deviation for a high quality oscillator; power law behavior is also common in this

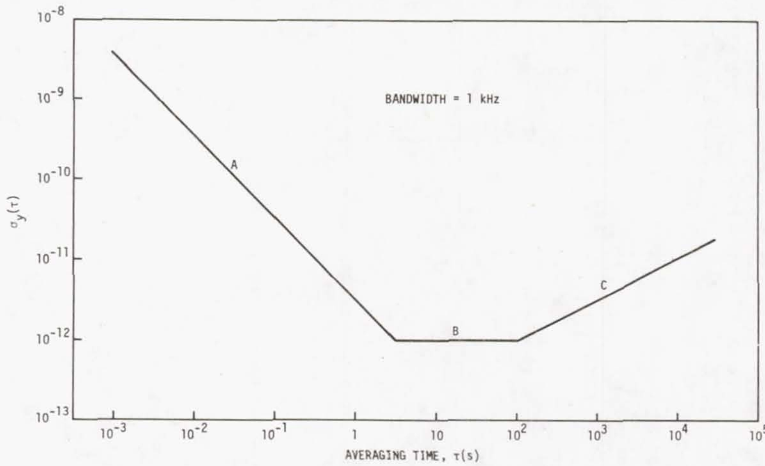


Fig. 9. Typical two sample deviation for a high quality oscillator.

case. In region A, $\sigma_y(\tau)$ is dominated by the high frequency noise of the oscillator and the stability usually improves as τ^{-1} (white phase noise) or $\tau^{-1/2}$ (white frequency noise). The stability is generally dependent on measurement bandwith in region A. In region B, the oscillator noise is dominated by the flicker of frequency behavior and σ_y is constant. For longer times the frequency stability generally degrades as $\tau^{1/2}$ or τ often due to deterministic effects like temperature, power level, and aging of components.

REFERENCES

[1] D. W. Allan and H. Hellwig, in Proc. Position Location & Navigation Symposium, (IEEE, New York, 1978) pp. 29-36.

[2] F. L. Walls and S. R. Stein, IEEE Trans. Instrum. Meas., IM-27, 249-252 (1978).

[3] S. R. Stein, C. M. Manney, Jr., F. L. Walls, James E. Gray, and Raymond J. Besson, in Proc. of the 32nd Annual Symposium on Frequency Control (EIA, Wash., DC, 1978) pp. 527-530.

[4] L. S. Cutler and S. L. Searle, Proc. IEEE 54, 136-154 (1966).

[5] J. L. Melsa and D. G. Schultz, Linear Control Systems (McGraw-Hill, New York, 1969) pp. 383-393.

[6] F. M. Gardner, Phaselock Techniques (John Wiley, New York, 1966) pp. 7-10.

[7] S. R. Stein and J. P. Turneaure, in Future Trends in Superconductive Electronics, Hugh C. Wolfe, ed., No. 44 (AIP Conference Proceedings, New York, 1978) p. 204.

[8] F. L. Walls in Proc. Eighth Annual Precise Time and Time Interval Applications and Planning Meeting (Technical Information & Administrative Support Division, Code 250, Goddard Space Flight Center, Greenbelt, MD 20071) p. 369-380.

[9] James A. Barnes, et al., IEEE Trans. Instrum. Meas. IM-20, 105-120 (1971).

QUESTIONS AND ANSWERS

MR. LAUREN RUEGER, Johns Hopkins University, Applied Physics Lab:

What servo systems have you suggested for use in complex systems and how would you propose to verify that they are functioning correctly? Would you have to go through a three-cornered comparison of three systems to verify they are working?

DR. STEIN:

You bring up a valid point. Of course, if one is talking about environmental problems, such as, say, acceleration sensitivity, one doesn't have a problem because one can compare the oscillator under acceleration to an oscillator in a nice environment, one that doesn't have a problem. When talking about pushing system performance to the limit, one always has the problem you discussed.

MR. RUEGER:

I believe what you are suggesting is an operational use of such a complex system; and if it is a complex one, then you need verification, in service, that these types of equipment are working and continue to work at their full performance.

DR. STEIN:

Perhaps I am not following your point. I think the answer to your question is that one tries, for instance, to produce an oscillator which has far better behavior, say, under irradiation than existing oscillators and still meets a time dispersion specification, long-term, something like a JTIBS-type specification.

And in that case, one doesn't really have a problem because one can compare the oscillator operating in the harsh environment to oscillators operating in good environments that maintain their good performance because they are not subjected to this problem.

MR. RUEGER:

The answer that you are giving me is one of the design verification, but not one in which the operation performance is sustained. And to verify that a device of this kind works in service over long periods of time is not the same as a qualification type of a test.

DR. STEIN:

I guess I will have to leave that to the systems designers.

DR. JACQUES VANIER, Laval University:

It may not be a servo system after all. You said it is a servo system, but you have to make measurements to calibrate everything around it, and that may be very expensive.

DR. STEIN:

Well, once again, the environmental measurements can and would be made by environmental sensors, automatically.

DR. VANIER:

Yes, but they are not in a servo loop. You're proposing an open loop.

DR. STEIN:

Well, it is open loop, and you are perfectly right in the sense that you are not attempting to control the environment. You are only attempting to optimize oscillator performance.

DR. GIOVANI BUSCA, Ebauches

I would like to ask if the good performance you have obtained depends on the fact that when you use your system, you correct for the phasing stability of the electronics. Is it essential to correct for the phasing stability of the electronics to have the best results?

DR. STEIN:

The answer to your question is really not trivial. The passive electronics may be an important part. We feel it is necessary. However, we have also used some novel crystals with new design, SC-cut, and this also is probably necessary.

Our results, for instance, with standard cut AT crystals, are approximately 3×10^{-13}, and I would guess that the reason for that has to do with temperature instabilities in the oven. So without the right crystal, you also cannot do it.

Page intentionally left blank

RELATIONSHIPS BETWEEN THE PERFORMANCE OF TIME/FREQUENCY STANDARDS AND NAVIGATION/COMMUNICATION SYSTEMS

H. Hellwig, S. R. Stein and F. L. Walls
Time & Frequency Division
National Bureau of Standards
325 Broadway, Boulder, CO 80303, USA

and

A. Kahan
Solid State Sciences Division
Rome Air Development Center
Hanscom Air Force Base
Bedford, MA 01731

ABSTRACT

Oscillators and clocks have moved beyond individual instrument applications and are now being used as critical elements of large scale electronic systems. The need for developing advanced time and frequency standards arises from increasingly more sophisticated DOD and civilian missions and functions in navigation, surveillance, identification and communication.

In these systems, time and frequency standards are employed to provide some desired performance or to be the central signal source for most systems functions. It is often not clear what frequency stability or accuracy is really needed. Often, a better clock or oscillator could simplify the system, or improve the performance of the system; in other cases, a different system design may lead to a reduction in the performance demands on the frequency or time standard.

This paper discusses the relationship between system performance and clock or oscillator performance. Our approach will be basic, pointing out some tradeoffs such as short term stability versus bandwidth requirements; frequency accuracy versus signal acquisition time; flicker of frequency and drift versus resynchronization time; frequency precision versus communications traffic volume; spectral purity versus bit error rate, and frequency standard stability versus frequency selection and adjustability. Our aim is to give the system designer and manager a better grasp of the benefits and tradeoffs of using precise frequency and time signals at various levels of precision and accuracy.

INTRODUCTION

During the last few decades, there has occurred a major revolution in the science and practice of time and frequency generation, dissemination and utilization. Frequency and time standards with very high performance characteristics have been developed, many of which are commercially available. Significant further progress in all important performance areas is possible. This includes improved short and long term frequency stability, spectral purity, size, weight, environmental insensitivity, cost, power demand and many more. In addition, there may be many cases where the application of a better clock could significantly relax other system parameters, and thereby reduce overall systems cost and complexity.

The question arises, what benefits can be derived from available or potentially available frequency standards and clocks? The PTTI Applications and Planning Meeting is the proper forum to address in a realistic manner this problem of long term technology planning. Our objective is to find a defensible rationale for long term investment strategy for time and frequency control device development. The usual approach to technology planning is to define system requirements and derive technology needs, which is followed by a program to achieve stated objectives. We have difficulties with this approach. Long term operational requirements are usually defined in very general terms, and it is very difficult to extract time and frequency generation and dissemination needs from stated goals.

We would like to suggest a possible alternate methodology. First, one should establish, for a particular application, the specific function performed by the device in the system and the elementary relationship between the time/frequency device and system performance. In addition, one should also determine the critical performance and operational time/frequency device parameters which limit system performance. After having established the quantitative relationship between the device and the system, the key question to consider is the system impact of improved time/frequency devices. How would an improved device parameter improve the system? It is hoped that by utilizing this approach, a systematic review of current and anticipated applications may point to a limited number of device parameters wich critically influence system performance. The optimum R&D investment strategy then is a program to improve these high system impact device parameters.

One possible methodology is to consider a matrix of (a) military or commercial functions and (b) evaluate the time/frequency device selection criteria. The functions consist of: navigation, positioning, and targeting; communications for command and control; surveillance and reconnaissance; intelligence; and identification. The selection criteria include: performance--stability, spectral purity, accuracy; system limitations--size, weight, power; cost--initial and/or life cycle; and

operational requirements. The operational factors consider: environment--space, nuclear, atmosphere; "ilities--reliability, durability, maintainability, interoperability; shock, acceleration, and vibration; warm up; and, finally, ease of deployment or man/machine interfaces. This paper is an initial attempt to define the qualitative relationship between clock performance and systems, and gives information which may be useful in deciding (a) how would better performing clocks improve system performance, or (b) what tradeoffs can be made between system performance requirements, the use of a better or worse frequency standard or clock, and other system design parameters. These questions are addressed from a purely technical viewpoint. It is beyond the scope of this paper as well as the competence of the authors to analyze system's tradeoffs and possible savings in system complexity and cost.

APPLICATIONS OF TIME/FREQUENCY STANDARDS

Overview

Application areas which most closely interact with frequency standards and clocks fall into the general categories of navigation and communication. Navigation has been linked for many centuries to clocks. Modern navigation ranges from the determination of the position of stationary objects on the surface of the Earth, to position finding of ships, aircraft and spacecraft, to some aspects of very long baseline interferometry. Navigation is characterized mostly by requirements on the long term stability, i.e., time keeping ability of clocks. Communication, which typically emphasizes the short term stability or spectral purity of frequency standards, ranges from exploitation of the electromagnetic spectrum to computer data links. Nevertheless, some forms of navigation, such as Doppler ranging, require also short term stability and spectral purity and some forms of communication, such as synchronous communication channels, may benefit from good phase or time stability of clocks. In addition, in a systems approach it may be beneficial to base systems which have both navigation and communication aspects on a common high performing time/frequency source.

Before we enter a more detailed discussion of any of these areas another important fact has to be noted. This is the ever present tendency to base specifications and systems performance requirements on the characteristics of a glamourous system component, in our case, on the frequency standard or clock. At the same time, there is often a lack of analysis whether the clock or other system parameters actually limit systems performance. In other words, a better clock may instantly lead to an improved system performance if other minor system limitations are removed, and/or serious system limitations would suggest the use of an inferior clock at no sacrifice in ultimate system performance. Similarly, procedural changes in system operations may be traded against

clock performance requirements resulting in benefits in system performance, complexity, reliability and operator dependence.

Tunability

Very frequently, the system designer is confronted with a need to provide for frequency or time (phase) adjustability. In the past, more often than not it appeared easiest to the system designer to specify an adjustable frequency standard or clock. Crystal devices can be made tunable by adding a capacitor in parallel or in series to the quartz crystal resonator. Tunability is then achieved by varying the value of this capacitor mechanically or electrically (varactor). Atomic frequency standards can be adjusted by changing the magnetic field which is applied to the interogation region of the atom (so called C-field region). The atomic resonance frequency has a second order dependence on the magnetic field in cesium, rubidium and hydrogen devices. This dependence is of the order of 10^{-4}/tesla (1 tesla = 10^4 gauss),[1] and rather large frequency changes can be affected.

A second method of providing frequency tunability is to use an adjustable frequency synthesizer in the electronic loop which is employed to servo the crystal oscillator (found in all atomic standards) to the atomic resonance. Since a certain frequency is needed to interrogate the atomic resonance frequency (here we assume a constant magnetic field), this variable synthesizer allows the generation of this fixed interogating frequency from a whole range of crystal oscillator frequencies, and thus provides tunability.

In most cases, however, it is overlooked that the addition of tunability affects the basic operation of the frequency standard or clock in a detrimental way. Frequency standards derive their high stability and accuracy from the fact that the essential control element, the quartz crystal resonator or the atomic resonator, has a high resonance Q, and that their resonance frequency is highly invariant with time or external parameter changes. The addition of a tuning capacitor, in the case of crystal oscillators, or the provision for magnetic field variability in the case of atomic resonators, virtually always degrades the performance by providing a direct coupling of the resonator to varying external influences. Similar arguments can be made for the variable synthesizer in atomic standards which affects the interogating microwave spectrum, and thereby again degrades the usable properties of the atomic resonance. The importance of this problem becomes clear if we consider high performance crystal or atomic resonators providing stabilities of the order of 10^{-13}. If the system designer requires tunability over a 10^{-7} region (often even larger is requested), this implies that we add an external influence capable of affecting the fundamental resonance phenomenon by 10^{-7} or more. If this approach is implemented, and at the same time the original 10^{-13} capability is expected, it means that the system designer expects all parameters operating on the tunability to be stable

to the 10^{-6} level. For example, if one volt at the varactor of a crystal oscillator provides the range of 10^{-7}, not only does the voltage have to be stable to one microvolt, but the capacitance versus voltage curve must be stable to the 10^{-6} level, and this stability has to hold, not only under idealized conditions, but also with respect to other parameters operating on these quantities, such as temperature, vibration, etc.

It is, therefore, useful to assert that large tunability and state-of-the-art performance in frequency standards and clocks are incompatible, and to the best of the authors' knowledge have never been satisfactorily combined. In other words, we have the axiom that tunability causes deterioration of clock and oscillator performance. Instead of requiring the clock manufacturer to supply a tunable, super-precision clock, system designers should consider implementing tunability by means external to the actual clock. To this end, frequency tunability can be achieved by adding an external direct synthesizer or a second tunable oscillator with a synthesizing loop. If phase or time adjustment is desired, external phase shifting by digital or analog means should be the method of choice.

Frequency synthesis is not without troubles of its own. The large amount of signal processing within general purpose synthesizers results in elevated levels of phase instability compared to the input signal. Consequently, the use of a synthesizer as the output source may prevent the attainment of the needed level of short and long term stability. A better solution is usually to use a (crystal) oscillator which is offset from the precision clock by means of a frequency synthesizer. This arrangement potentially provides much improved spectral purity and long term stability than the direct use of a synthesizer as the output device; however, the frequency tunability is restricted in speed and range.

Warm-up and Environmental Sensitivity

From the point of view of the user, especially in military applications, some usually neglected properties of frequency standards may be the most important. These include the behavior in severe environments and the time required to achieve a specified accuracy and stability after a cold start. Such properties are usually not optimized by the design which yields the best long term stability or spectral purity.

Practical or engineering problems, including the need for ovens and vacuum pumps, limit the warm-up performance of both laboratory and commercial frequency standards. However, there is one device which was designed for fast warm-up, the passive ammonia frequency standard.[2] This device achieves its full accuracy of 10^{-9} shortly after turn-on from a cold start. Despite the unimpressive accuracy compared to com-

mercial cesium standards, the overall performance of such a device for short duration missions may be significantly better.

New schemes for temperature compensation of quartz crystals may also result in rapid achievement of 10^{-9} or better accuracy from a cold start. In addition, the application of "smart servos" may allow several oscillators to be combined in ways which improve the overall environmental insensitivity. Improvements may be made in the performance not only with regard to turn-on, but also in severe acceleration and radiation environments. Such servos would have both variable gain and time constants. This concept amounts to a systems approach to the frequency standard itself.

Synchronization

The effect of stability, accuracy and the ability to synchronize clocks on the performance of a complex system is well illustrated by the following sample problem: Find position with respect to a natural or artificial geographic grid, such as the surface of the Earth, or a network of orbiting satellites. An ideal system of this nature is global in coverage, three-dimensional, arbitrarily precise, providing instantaneous access and position fix, and able to accept users automatically and instantaneously. Another example for resynchronization requirements is a secure communications system with access and identification being provided via timing.

Of primary importance for the functioning of a time-ordered system of this nature, is time accuracy. Here accuracy is meant in the sense of time deviation of individual system clocks from system time, and user's knowledge of system time. In navigation, position information often is extracted from time information using the propagation speed of electromagnetic signals (speed of light); therefore, position accuracy of one meter requires a time accuracy of 3 ns.

In principle, system synchronism can be assured arbitrarily well via electromagnetic signals, providing time information to and between the components of the system. On the other hand, requirements of reliability, freedom from effects which deteriorate signal transmissions (natural and artificial interference, weather conditions) and autonomy of subsystems (ability to operate in case of partial system failure) would lead the designer to a system requirement of no reliance at all on electromagnetic signals. In other words, these latter requirements make ideal clocks desirable, having, once synchronized, no time deviation. Real systems, of course, fall in between those two extremes featuring finite resynchronization times. Nevertheless, the general statement can be made that a more stable clock would allow the relaxation of resynchronization requirements, and correspondingly increase reliability, autonomy and interference immunity.

For most clocks and oscillators one may write as a useful description of the time deviation [3]

$$x(t) = x_0 + y_0 t + 1/2Dt^2 + \varepsilon(t) \tag{1}$$

The first three coefficients in this equation are model parameters: $t = 0$ signifies the instant of synchronization, x_0 is an estimate of the synchronization offset at $t=0$, y_0 is an estimate of the fractional frequency offset at $t=0$, and D is the estimate of the fractional frequency drift. The fourth term, $\varepsilon(t)$, represents the random frequency fluctuations of the oscillator which are characterized by the variance:

$$\sigma_y^2(\tau) = \left\langle \frac{(y_{i+1}-y_i)^2}{2} \right\rangle \tag{2}$$

where y_i denotes the fractional frequency measurement of duration τ and $\langle\rangle$ denotes time average. All modern clocks and oscillators have been found to typically improve in stability with increased sampling times (region 1) until they reach their stability limit or flicker floor (region 2). For long sampling times τ, they finally deteriorate in stability (region 3). The errors in the model parameters and the frequency fluctuations all contribute to the time deviation error. Table I summarizes these three different regions and two model parameters and gives

Table I.

	Frequency fluctuation	Time deviation $x(t)$
Region 1	$\sigma_y(\tau)=k_p\tau^{-1}$	$k_p/\sqrt{3}$
	or	
	$\sigma_y(\tau)=k_f\tau^{-\frac{1}{2}}$	$k_f t^{+\frac{1}{2}}$
Region 2	$\sigma_y(\tau)=k_F\tau^0$	$\frac{1}{\sqrt{\ln 2}} k_F \cdot t$
Region 3	$\sigma_y(\tau)=k_w\tau^{+\frac{1}{2}}$	$k_w \cdot t^{+3/2}$
Offset	y_o	$y_o \cdot t$
Drift	D	$\frac{1}{2}Dt^2$

quantitatively the corresponding time deviation, x(t). It should be noted that the time deviation given in each of the five rows of table I assumes that the corresponding frequency fluctuation is the sole effect for all averaging times.[4] A more complete treatment of this behavior is given in reference 5. However, table I serves as an adequate quantitative guide because for any particular range of t the dominant effect can be easily identified by comparison of the relative values from regions 1, 2, 3. In such treatment the various coefficients k also become model parameters.

We should first examine table I. If we assume that we typically deal with relatively long times between resynchronization we can ignore the aspects of region 1. Region 2 and the frequency offset have in common that time deviation is proportional to elapsed time. Therefore, figure 1 depicts the conditions of frequency offset and/or flicker of frequency. Figure 2 depicts conditions where random walk of frequency (Region 3) dominates. In figure 3 the condition of a clock with frequency drift is shown. In the three figures, the fractional frequency offset or flicker floor (fig. 1), the frequency walk coefficient (fig. 2), and the frequency drift coefficient (fig. 3) are plotted versus the time between resynchronization; the desired or required timing accuracy (position accuracy) is the parameter.

Instantaneous access is the other system parameter which is related to clock performance. An example is best used to illustrate this. If one meter navigation accuracy is the system design goal, then a 3 ns capability must be available not only from the system but also from the user clock. The minimum requirement, therefore, on both clocks is x(t)=3 ns independent of time, or $\sigma_y(\tau) = k\tau^{-1}$ with k = 3 ns. This is a requirement which most clocks will fulfill for sampling times from less than a second to many seconds. If we want to require that the user has an instantaneous position fix entering the system or has no means to acquire system time then the user must have an a priori knowledge of system time to three nanoseconds. If the user had an initial, perfect knowledge of system time, then goes on a mission, (being cut off from direct access to system time), a clock with certain performance would allow him to have a mission duration of one hour while retaining a one meter positioning accuracy, whereas a clock with better performance in regions 1 and/or 2 would extend this to, for example, many days.

Another interesting aspect in this context is the tradeoff between signal-to-noise, (antenna design, signal strength) and clock performance. If, for example, system time can be acquired by rapid scanning on a second-by-second basis, across four satellites (the GPS system) then one nanosecond timing is required only for the duration of a few seconds; thus, a relatively inferior oscillator would be sufficient. If, however, more elaborate antennas and pointing are required and an acquisition time per satellite of, for example, several hundred seconds is needed, then one nanosecond needs to be maintained over one thousand or

more seconds requiring a very good oscillator performance on the users' side, as can be seen from figures 1-3. Related arguments can be made for geodetic application of very long baseline interferometry, or for the VLBI applications in deep space tracking.

An interesting example of the relationship between long term stability and recalibration interval comes from the area of communications. In order to increase the available spectrum, operating (carrier) frequencies are being increased while the channel width is held nearly constant, resulting in tighter specification on the long term stability of both transmitter and receiver frequency references. In the case of low cost mobile transmitters, the present specifications are just met by the crystals in use. In order to accomplish the next step higher in frequency it will be necessary to either improve crystal performance, ovenize the crystal, or use a transmitter/receiver design which calibrates itself using the received signal from a more accurate base station. If the last alternative is adopted, the mobile transmitter frequency standard may exceed specification limits between uses. This situation could be corrected by requiring that after the calibration interval such a mobile radio receives a transmission from a base station before transmitting itself.

Signal Aquisition

Systems which deal with very weak signals are sensitive to the spectral purity as well as the long term stability of the system oscillators. In this portion of our paper we illustrate the effects of stability, spectral purity and signal-to-noise ratio with examples from Doppler radar and Doppler navigation.

The simplest radars are CW Doppler devices which determine the radial velocity of a target by homodyne detection of the return signal with a transmitted signal. Phase noise in the vicinity of the frequency of the carrier occurs, of course, in the same small frequency range around the carrier as the Doppler shifted return signal from the target. Therefore, the relatively weak reflections of the phase noise from nearby objects with a very large cross section cannot be distinguished from the Doppler shifted signal of interest which, though initially strong, comes back at very small intensity from the distant target having small cross section. The level of phase noise, therefore, relates to signal contrast and a decrease in phase noise will extend the range and/or the contrast of this Doppler system.

Present day navigation systems for satellites and for deep space probe tracking utilize round trip coherent measurements. The signal is transmitted to the spacecraft where it is transponded for Doppler detection on the ground. The frequency instability of the oscillator during the round trip time of the signal to the spacecraft and back is proportional to the velocity error thereby establishing a requirement on the medium

term time stability of the frequency standard. Figures 1-3 can be used to determine the quantitative requirements. This two-way Doppler tracking has the disadvantage that the spacecraft must detect the transmitted signal with a very small antenna which requires large and accurate transmitting antennas with associated expenses. If an onboard frequency standard is included into the spacecraft the size and complexity of the Earth-bound antenna can be highly reduced. The spacecraft beacon is then measured and compared by two separated ground stations, in a way analogous to very long baseline interferometry. In both cases, position and velocity accuracy depend directly on the synchronization and syntonization of the clocks at the two ground stations. If 10 cm range error and .05 rad angular error are tolerable, then subnanosecond timing must be realized at the two groundstations.

A receiver for coherent communications must reconstruct the frequency and phase of the unmodulated transmitter carrier. This is normally accomplished by phase-locking a voltage controlled oscillator (VCO) to the received signal. The speed and accuracy with which this is done depend on the stability and accuracy of the signal as transmitted, the stability and accuracy of the VCO, and the signal-to-noise ratio of the received signal.

If the frequency offset, Δf, between the transmitted signal and the VCO lies within the phase-lock loop (PLL) bandwidth, B_L, then lock is acquired within one rf cycle. However, the PLL bandwidth is limited by the requirement that within this bandwidth the weak received signal must have signal-to-noise ratio greater than one. Thus, the usual situation is one where Δf greatly exceeds B_L. For a critically damped, high gain, second order PLL, the pull-in time is approximately[6]

$$T_p = \frac{4}{B_L}\left(\frac{\Delta f}{B_L}\right)^2 \tag{3}$$

In the case, for example, of a loop with $\Delta f = 100$ Hz and $B_L = 1$ Hz, the pull-in time would be more than 10 hours.

The result of this situation is that acquisition is generally achieved by sweeping the VCO. The maximum rate is limited by the bandwidth of the loop. Consequently, the pull-in time is approximately

$$T_p = \frac{4}{B_L} \cdot \frac{\Delta f}{B_L} \tag{4}$$

which would be only 0.1 hours for the above example. There are, of course, other more complex PLL and PLL/frequency lock loops, some with variable sweep characteristics, which can reduce the pull-in time for large initial offsets.

Once lock is acquired, the VCO phase reflects the transmitted carrier phase to a degree determined by the transmitted carrier phase noise, the VCO noise and the noise of the receiver. The residual noise results in errors in transmitted data and limits the data rate.

Thus, improvements in the a priori knowledge (accuracy) of the VCO would correspondingly reduce the PLL acquisition time via reducing the pull-in delay and/or the needed sweep/search amplitude or time. In addition, lower phase noise of the VCO allows faster data rates and/or lowers the error incidence.

OUTLOOK

The above discussion shows that navigation and communication systems can benefit if better frequency standards and clocks become available. "Better" means not only a further increase in fundamental stability and accuracy, but, even more importantly, availability of adequate stability and accuracy at attractive cost, size, power demands, as well as environmental insensitivity. For example, many applications would become commonplace as well as economical if a user-oriented frequency standard which provides 1 ns timing capability from less than one second out to several thousand seconds, and which performs under severe environmental conditions would become available. Such a standard still does not exist. Therefore, development has to focus on both the area of subnanosecond timing out to many days in a practical, rugged, environmentally insensitive device at reasonable cost, as well as on the area of user-oriented standards. In addition, oscillators of low phase noise which are practical and environmentally stable are needed.[7]

REFERENCES

1. Helmut Hellwig. Atomic Frequency Standards: A Survey, Proc. IEEE, 63(2):212-229, 1975. Also in Proc. 28th Ann. Symp. on Freq. Control, Ft. Monmouth, NJ, May 1974, p. 315-340.

2. David J. Wineland, David A. Howe and Helmut Hellwig. Special Purpose Ammonia Frequency Standard, Proc. 8th Ann. Precise Time and Time Interval (PTTI) Planning Meeting, Dec. 1976, p. 429-448. Also in IEEE, Trans. on Inst. and Meas., Vol. IM-27, No. 1, March 1978.

3. Helmut Hellwig, David W. Allan and Fred L. Walls. "Time and Frequency", Atomic Masses and Fundamental Constants, J. H. Sanders and A. H. Wapstra, eds., p. 305-311 (Plenum Press, 1975).

4. Jacques Rutman. Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress, Proc. IEEE, 66(9):1048-1075, 1978.

5. David W. Allan and Helmut Hellwig. Time and Phase Dispersion of Clocks and Oscillators, Proc. IEEE Position Location and Navigation Symposium, San Diego, CA, November 7-9, 1978.

6. F. M. Gardner. Phaselock Techniques (J. Wiley, New York 1966).

7. A wealth of relevant and important information can be found in G. M. R. Winkler's "Timekeeping and its Applications," Adv. in Electronics and Electron Phys. 44, 33-97 (1977) (Acad. Press).

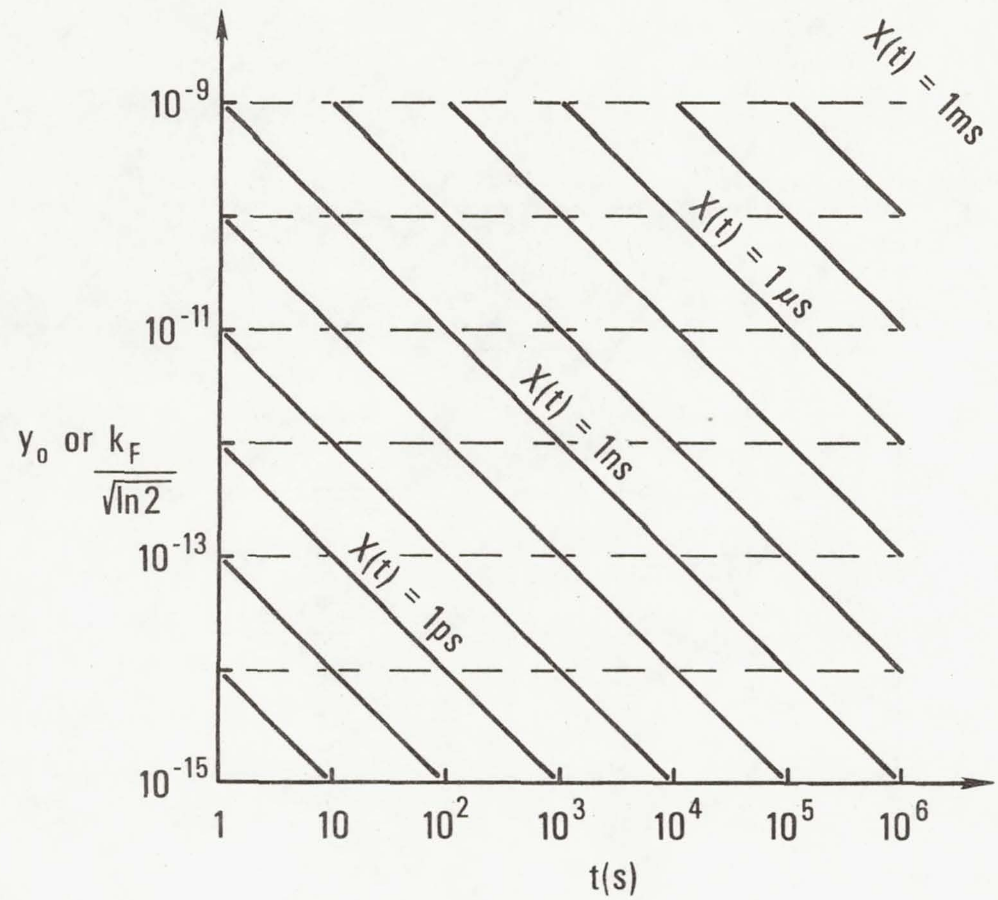


Figure 1. Largest allowable frequency offset, y_0, and flicker of frequency floor, k_F, as a function of time, t, after perfect synchronization. The achievable time accuracy, X(t), is the parameter.

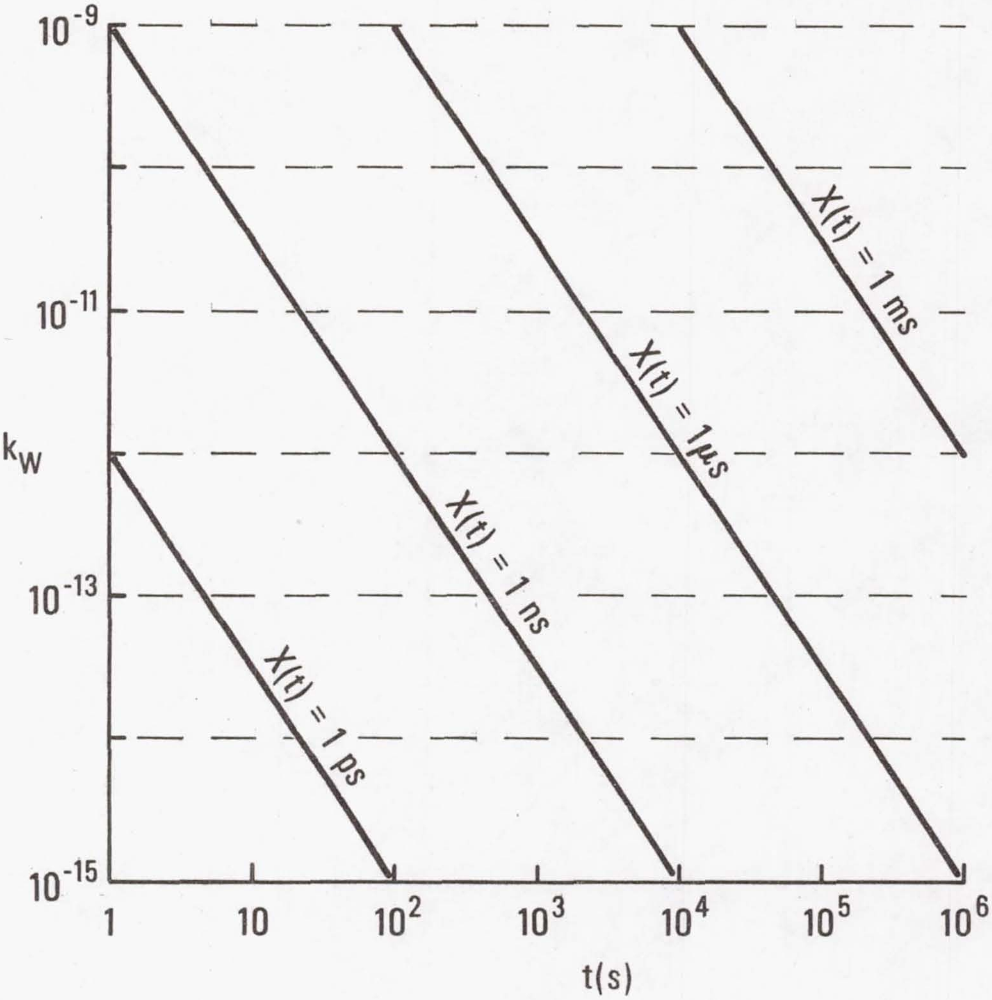


Figure 2. Largest allowable walk of frequency coefficient, k_W, as a function of time, t, after perfect synchronization. The achievable time accuracy, X(t), is the parameter.

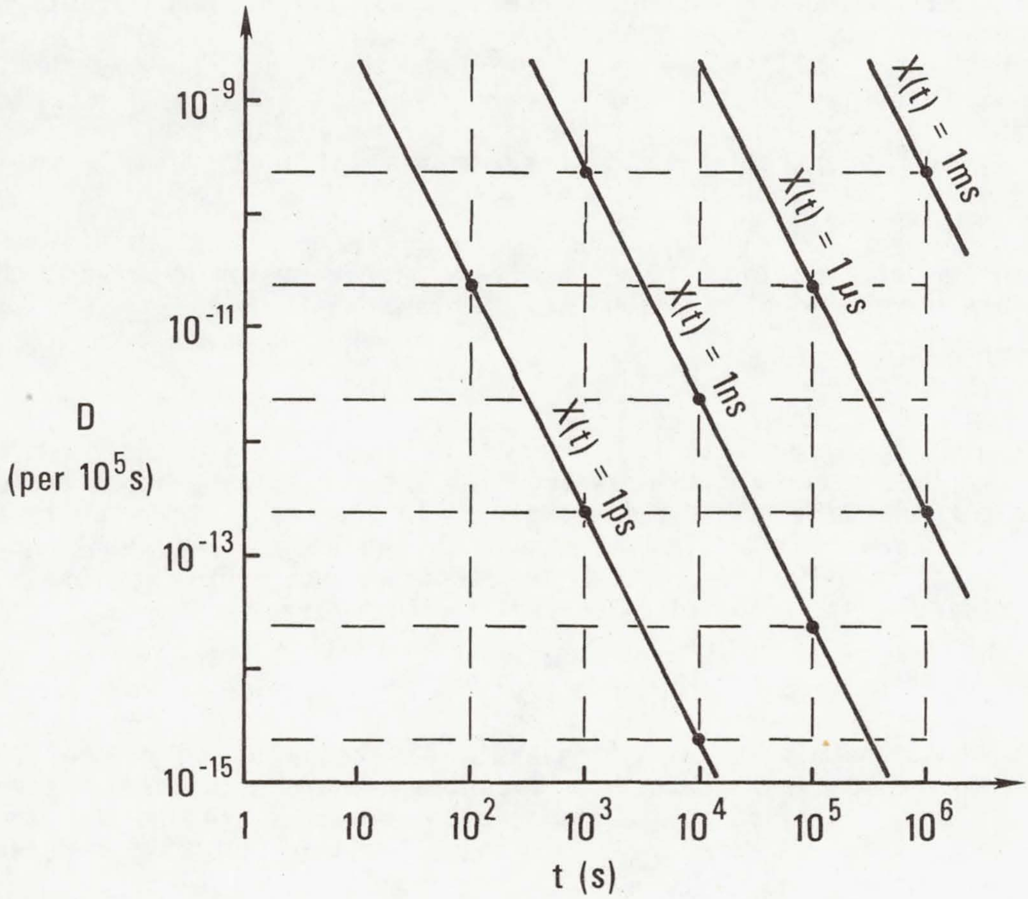


Figure 3. Largest allowable frequency drift coefficient, D, as a function of time, t, after perfect synchronization. The achievable time accuracy, X(t), is the parameter. D is used as a fractional frequency drift per 10^5 sec (approx. 1 day).

QUESTIONS AND ANSWERS

MR. HARRY PETERS, Sigma Tau:

I just wanted to comment on the tuneability equation relationship between range of tuneability and the stability of the oscillator. I think you will agree with me that there are now synthesizers which require on the order of half a watt, no phase lock loops, and give, for example, for the hydrogen maser, 10^{-16} or 10^{-17} range, digitally controlled, and where that relationship really wouldn't hold.

These are already developed and are operating in two laboratories at this time. They are very simple system-wise and, therefore, probably not a reliability problem.

DR. HELLWIG:

I totally agree with you, Harry. My answer to that is that it is sometimes difficult to draw the line between talking about the basic clock and talking about the system. In your case (in my philosophy, if you want) the line is drawn at the maser. And what you are saying, really, is what I am recommending: Put the tuneability outside of the basic clock concept. I fully agree with you.

MR. SAM WARD, Jet Propulsion Laboratory

I don't feel that enough emphasis has been put on reliability in these approaches because not even the reciprocal nature of time and frequency is effective if you can't keep the system going. And redundancy is not the answer, either, because that only complicates the problem.

DR. HELLWIG:

Do you mean reliability of the clocks themselves or reliability of the system and clocks?

MR. WARD:

Basically, the clocks themselves.

DR. HELLWIG:

I deliberately stayed away from that aspect and I think we will stay away from it in the printed paper. I think the only thing I can say

is to repeat what Dr. Winkler said already: If you plan to run operational systems, please put in clocks which have proven themselves.

DR. GERNOT M. R. WINKLER, U.S. Naval Observatory

And have redundancy. I will put in my five cents worth also, and offer a conjecture to Dr. Hellwig's very excellent observation that, unfortunately, the manufacturers have not been turned on by, let's say, the intermediate range because they know that only the best will be specified. But the situation is opposite for the OEM market for understandable reasons. Only the least you can get away with will be specified.

So I think the situation will change as the OEM market expands, and more and more of these intermediate performances will be required, for instance, in counters and other such things. These have been the major driving force for the development of very nice little crystal modules which are available today at relatively low cost.

Page intentionally left blank

ACCELERATION SENSITIVITY COMPENSATION IN HIGH PERFORMANCE CRYSTAL OSCILLATORS

Donald A. Emmons
Frequency & Time Systems, Inc.
Danvers, Massachusetts

ABSTRACT

High stability crystal oscillators are an essential component in a wide variety of systems, from satellite based navigation and communications, to Doppler radar, time keeping and so on. In a hostile shock and vibration environment, the full potential of such systems is severely limited by the intrinsic acceleration sensitivity of the widely used resonator types. Thus, it becomes increasingly important to investigate the possibility of increased frequency stability under adverse conditions.

Two approaches to achieving reduced sensitivity will be discussed. The first involves electronic compensation within the frequency control loop. The second utilizes two resonators of comparable acceleration sensitivity to compensate each other.

Problems encountered in matching and tuning the resonators will be discussed, as well as orientation symmetry of the frequency deviation patterns. Results on frequency stability which reflect an improved static sensitivity of less than 5 x 10^{-11} per g are presented.

INTRODUCTION

High performance crystal oscillators are critical components in navigation, communications and time-keeping systems which require the best attainable frequency stability and spectral purity. This high stability performance is usually realized only under quiet ambient conditions, yet systems are often required to operate, without degraded performance, in a shock and vibration environment.

Although new designs (Ref. 1) and improved resonator types will ultimately be available, the most widely used type at present is the AT-cut, which shows a typical acceleration sensitivity of 10^{-9} per g or greater. Some recent progress in compensating AT-cut overtone crystals for reduced sensitivity to low frequency mechanical inputs will be described. This is a continuation of the work briefly mentioned previously at these meetings (Ref. 2). The basic test bed has been the rugged high performance Model 1000 oscillator using an oven stabilized, 5th overtone 5 MHz resonator.

First, the observable effects associated with shock and vibration will be discussed and correlated with the static g sensitivity of the resonators. Particular attention is given to the spatial dependence of frequency change with relative acceleration direction, and the symmetry of this pattern.

Two basic approaches (neither of them new) have been explored and each has some limitations. The first scheme encompasses various methods of correcting the VCXO tuning via an accelerometer derived signal. Results from a straight forward systems approach using an external sensor have been reported recently by Przyjemski (Ref. 3) We have achieved similar improvement using different resonator types. The disadvantage of using a VCXO whose tuning is non-linear will be pointed out.

In the present work we have also incorporated sensors within the crystal oven for high stability, compactness, and potentially better high frequency compensation response (Ref. 4).

The alternate approach has been to use two resonators in series and arranged so that acceleration induced frequency shifts cancel as suggested by Gagnepain and Walls (Ref. 5). Here the individual g sensitivity pattern symmetries are extremely important. Several interesting problems have come to light. Among the advantages on the other hand are that resonator and sensor have nearly identical construction, and they can be placed physically close to one another. Furthermore, a very interesting experimental effect linking temperature induced stress and g-sensitivity has been found, which has to do with thermal shock and warm-up behavior in AT-cut quartz.

MOTIVATION

Continued effort toward improving the performance of existing resonator types is justified by the immediacy of demands on existing technology, and partly by the fact that it can be done.

Although considerable data have been gathered on a variety of resonator types, we concentrate here on 5th overtone oscillators having particularly good signal-to-noise ratio close in to the carrier, and short-term stability performance of better than 10^{-12} in the 1 - 100 second region. (See Figure 2)

Low noise performance becomes quickly degraded in a hostile physical environment. At a vibration frequency f_m, modulation sidebands appear with a magnitude $(\gamma a f_o/2f_m)^2$ where γ is the intrinsic g-sensitivity coefficient, a, the peak acceleration, and f_o the carrier frequency. (See Figure 3)

For systems involving frequency multiplication to high order, n, noise appearing close to the carrier and multiplied up by n^2 is of concern. Since the

sideband level goes as f_m^{-2}, perturbations down to zero frequency dominate, while with increasing frequency both electronic and mechanical filtering are relatively easy and the sideband level is usually decreasing.

Acceleration inputs may span a wide range of amplitude, and be randomly oriented. The spectrum extends to d.c. where attitude in the earth's gravitational field may be a slowly varying parameter. In a closed loop system, accumulated phase error may be of interest (Ref. 3).

In frequency standards, such as the cesium beam instrument, frequency slewing of the flywheel oscillator may occur at a rate such as to give unacceptable frequency offset, where the time rate of oscillator drift multiplied by the control loop time constant gives the resultant error. This is particularly severe when τ has been made long to exhibit short-term stability approaching the performance of the open loop oscillator. Notice, however that phase slip resulting from frequency shift in the oscillator can be reversible, averaging to zero if the mean value of g-induced shifts is zero.

Another effect can occur under sustained vibration. At the onset of 20 Hz 1 g vibration, we have observed a transient shift of the order 5×10^{-10} recovering in 10 minutes to within about 1×10^{-10}. At the cessation of vibration, the opposite transient occurs, and the frequency settles back to nearly the original. This effect is traceable to perturbed oven control, which can be minimized by packaging techniques.

The effect of mechanical shock on high precision crystals is generally seen as a frequency jump either positive or negative. For 2300 g pyrotechnic shocks, jumps are typically several parts in 10^{10} per pulse, with many successive shocks leaving the final valve within 5×10^{-10} of the original frequency.

Shock mounting of the crystal enclosure itself is possible. However, for effective compensation it is necessary to have the resonator unit rigidly coupled to the compensating element, so as to avoid phase differences between the two responses. It is thus, more useful to shock mount the oscillator as a whole to attenuate high frequency inputs, even in cases where only low frequency vibration is expected.

RESONATOR BEHAVIOR

The response of AT-cut quartz to acceleration stresses has been discussed by many authors. In relation to the resonator geometry it is found that the axis of maximum acceleration induced frequency change is dependent on, among other factors, the angle of cut, and position of resonator supports with respect to the crystal axes. Valdois and Besson (Ref. 6), Lee (Ref. 7) and others have discussed the expected stresses and it has been shown that Δf along a given axis is linear with g to at least 50 g. In general an approximately cos Θ dependence for the force-frequency effect is found, implying linear response to the vector component along a principal axis. Our results with high Q resonators bear this out.

Given that linearity and regular angle dependence exist, one can then make detailed measurements of spatial symmetry at a level of 1 g. Figure 4 shows the pattern for 2π rotation in azimuth. The plano-convex quartz disk is supported in a three point ribbon mount. The acceleration of 1 g is fixed, downward, and the resonator rotates clockwise giving the associated values of $\Delta f/f$ shown. A coefficient γ' for maximum Δf in this plane is 4×10^{-10} per g. The heavy circle marks the reference line of zero deviation. It should be emphasized that these are simply one set of data, for a particular resonator, in one special plane. Experience shows that the correlation of symmetry pattern to the position of the mounting supports is of secondary interest. More important is the high degree of

symmetry, indicating that in this plane a single axis accelerometer could give exact cancellation.

The appearance of the plot is a function of purely arbitrary scale factors. The expression is $A + \gamma' \cos \Theta$ and for γ' small compared to A, the data lie on a displaced circle.

The same resonator is shown in Figure 5, for rotations in polar angle Θ, about an axis through the odd support point and the plane of the disk. Note that γ is 1.3×10^{-9} per g and that the symmetry axis is tipped slightly from the vertical. (Data for 3rd overtone resonators show a similar shape with coefficients as much as 2 times greater.) The data points shown lying nearly on the zero circle illustrate the compensation which was achieved with a simple accelerometer incorporated within the oscillator, to be described in the next section. The exact degree of cancellation depends on fine tuning of axis angles and compensation magnitudes.

This is similar to what Przyjemski has already shown for an oscillator compensated by an external accelerometer.

To introduce the discussion of compensation schemes, Figure 6 shows some points which need to be considered for ideal performance, recognizing that non-ideal performance will also be useful.

1. Perfect symmetry is required so that no residual sensitivity (to the desired low level) is observed. This can also be taken to mean that the coordinate systems of resonator and compensator must be exactly coincident in space. Otherwise, angular rotation about an axis other than the line of centers produces centripetal acceleration which is distinctly not compensated, but shows a coefficient 2γ.

2. Linear frequency tuning may be important for a system which uses a compensation signal to control the VCXO. At any given tuning voltage the frequency deviations are small enough that the linearity is sufficient even for typical varicap tuning; however, as the oscillator ages, the desired correction coefficient may vary drastically.

3. Vibration induced sidebands fall off as f_m^{-2} (typically -120dB for 1 g at 10 KHz). Thus, large compensation bandwidths may not be necessary. An upper limit of 1 KHz might be typical, at which point mechanical isolation is easy. In any case, above several KHz one expects trouble from phase shift due to the resonator ribbon mount resonances.

   It is interesting to note, however, that oscillators can exhibit an fm bandwidth of up to 20 KHz when the usual low-pass VCO filter is removed.

4. All spurious sources of phase modulation must be suppressed before the intrinsic quartz effect can be linearly compensated.

5. Vibration induced frequency shifts can be complicated by the exagerated shifts from "thermal shock" accompanying very small crystal oven temperature perturbations.

6. Hysteresis includes effects which can not be compensated, such as shock induced permanent offsets.

7. Finally, one wishes to compensate in such a way as to introduce no additional noise into the oscillator loop, and no additional sources of apparent aging.

Relative to point 2 above, an FTS oscillator whose linear tuning range was extended to accomodate a particular g-compensation sensitivity, is shown in Figure 7 to have integral non-linearity of less than 5%.

COMPENSATION SCHEMES

We now turn to some considerations of the series resonant circuit containing a high Q thickness shear mode crystal. The frequency determining network is such that a series (capacitive) reactance of 1 Kohm typically pulls the crystal to the operating frequency. The tuning rate dX/df is about 100 ohms per Hz. Thus, given a g-sensitivity of 10^{-9} per g, the compensation required is df $\cong$ 5 milliHz per g, or dX $\cong$ 0.5 ohm per g. Figure 8 shows the circuit. The parallel combination of reactances (usually capacitative) determines the resonant frequency. This approximate analysis will ignore the resonator electrode capacitance and lead strays, lumped together as X_o'. X_1 is a calibration capacitor on the order of 1 Kohm at 5 MHz. X_2 is a variable reactance which provides electronic tuning. X_3 indicates various possible stray capacities adding to X_o'. Supposing X_3 to be quite large (stray C small) we have that the effective series X is just $(X_1 + X_2)$, controlling the frequency at 100 ohms per Hz.

The intrinsic resonant frequency shifts of interest are of order 5 mHz per g. If X_1 or X_2 were to be mechanically or electrically modulated with a coefficient 0.5 ohm per g, then frequency compensation would be achieved. With X equal to 1 K (32 pF) we need a capacitive change of 15 milli-pF per g; or, if 1 K inductive, then about 15 nHy. This, with variations, encompasses one class of compensators. One either modulates the reactance X_1 gravitationally, or modulates X_2 indirectly from an external transducer.

Now consider the case where X_3 is not negligible. Because stray capacitance to ground is often several pF, it is easy to imagine mechanically modulated changes of the order of milli-pFs (femto-Farads). Suppose a 1 pF stray were due to leads or components spaced 0.1 cm from ground. Then for spurious modulation to be less than 1 mpF, the spacing must stay constant to 10^{-4} cm. This means that leads and components must be firmly fixed. In fact, the conditions are less stringent since X_3 is shunted by much larger capacitances.

However, this leads directly to the idea of mechanically modulating X_3 in the same manner as for X_1.

A generalized capacitive transducer which responds linearly to acceleration (deflection proporational to force) will have a fractional sensitivity $\Delta X_3/X_3$ per g which is sufficiently linear for small deflections. Frequency response down to d.c. is of course assumed. As an example, changes of 10 mpF per g or so can produce the desired half-ohm compensation signals. (We have, with apologies to Michael Faraday, a 'milli-puffer').

The overall sensitivity in the network goes as $\Delta X_3/X_3)^2$ so that X_3 must be kept small if we wish to avoid diluting the effect. The device reactance will be proportional to some spacing, s, and fractional change in X will be proportional to fractional change in s. Acceleration induced deflections are fixed by material stiffness and mass distribution. Fundamental considerations of the dynamic response show that a natural resonance frequency of 500 Hz is reasonable. Experimental devices tend to show lower frequencies than this, but the technology of capacitance microphones has pushed response well up above 1 KHz for such devices.

Returning to the expression $\Delta X \cong (X_1 + X_2)^2 \Delta X_3/(X_3)^2$ we note that the shunt effect of X_1 and X_2 dilutes the transducer coefficient unless X_3 is kept small. However, X_3 cannot be too small because that dilutes the oscillator tunability. Some compromise must be chosen.

It is also clear that the compensator coefficient is a function of X_2 and so depends on oscillator tuning; the effective compensation may change by 20% over a tuning range of 2 Hz. This is unfortunate for the milli-puffer. However, an oscillator intended for fixed frequency operation can be exactly compensated. Quartz aging would then require eventual readjustment.

A question which should be raised is whether the acceleration induced frequency shift intrinsic to the quartz gives an observed Δf which depends on external circuit parameters? Empirically the answer is no. Over a 2 Hz range the observed sensitivity $\Delta f/f$ per g is constant to within $\pm 2 \times 10^{-11}$. The required compensation is apparently not a function of operating point on the piezoelectric reactance curve.

Figure 9 shows results of reduced g-sensitivity using external control of the parameter X_2. This oscillator has been subjected to vibration testing in the range 8 to 33 Hz, with maximum g level up to 1.5 g, and the measured phase noise was -98 dB (15Hz) relative to the carrier, which checks quite well with the measured static sensitivity. The results that were shown in Figure 5 were obtained using the X_3 'milli-puffer' method.

The second type of g-compensation uses two resonators in series (Ref. 5). The method is indicated in Figure 10. Initial success with this method is indicated by the low phase noise. The dual crystal within the oscillator loop exhibits the same loaded Q as for a single resonator, at the same time showing reduced g-sensitivity. Results were disappointing however, in that the nearly exact cancellation expected did not occur. What happens is that in pairing the two resonators, the relative stray capacitance for each is changed, and the individual sensitivities no longer cancel.

The feature around 2.5 Hz is associated with an incipient frequency instability which grows worse (and shifts in frequency) as the relative strays are

adjusted to achieve exact compensation. In the analysis of the dual resonator equivalent circuit, two allowed frequencies appear (Ref. 5). These may be widely separated and yet permit some degree of g-compensation. What is surprising is that the compensation can be 'tuned' at the expense of overall stability. At exact compensation a slow beat appears, correlated with the phase noise anomaly. This behavior is still under investigation.

The selection process for matching resonators is indicated in Figure 11. Here the two sensitivities are nearly identical aside from a small relative tilt. Following a parity reversal of one set of geometric axes, the two resonators are put in series and installed in the same oscillator.

Figure 12 shows results for a pair of 3rd overtone crystals. The sensitivity for one alone is 19×10^{-10} per g and for the combination, is reduced to 3.5×10^{-10} per g.

An interesting consequence of the partial cancellation of stress induced frequency shifts is found in the thermal shock behavior of the dual crystal units. The warm-up behavior is shown in Figure 13, along with the warm-up curve when one of the resonators is bypassed. The latter shows the typical 0.5 Hz to 1 Hz overshoot due to 'thermal shock' non-equilibrium. It is not surprising to find that frequency shifts induced by thermal gradients are intimately related to those induced by mechanical stress; what may be surprising is that in a particular oven geometry these stresses can be compensated to any extent during the warm-up time. This is reminiscent of early attempts (Ref. 8) to reduce the equilibrium temperature coefficient in A element resonators by operating two in series. This new phenomenon may shed some light on the non-time-equilibrium conditions.

SUMMARY

Within the frequency determining network for a high precision crystal we have a variety of possibilities for making use of acceleration derived compensating signals. For a capacitive device, sensitivity of the order of tens of femto Farads per g is needed, and frequency response down to zero is desired. The analysis is straight forward in terms of series reactances. Inductive g-sensitive transducers may also be employed.

Furthermore the reactance of interest may also be controlled elctronically by an externally derived correction signal. The bandwidth of the oscillator sustaining loop is more than adequate for handling the desired compensation.

There seems to be no clear-cut answer as to which scheme is preferable in all cases. All of the single crystal methods have had comparable success, and the dynamic tests have been consistent with measured reductions in static g-sensitivity.

The dual crystal measurements have produced evidence of a fundamental frequency instability, but compromises can be made for less than complete compensation, with overall improvement. Of particular interest is the fact that thermal shock frequency deviations, as noted during warmup, are reduced from the usual single crystal case.

In conclusion, Figure 14 shows some comparisons of compensated vs. uncompensated performance. The example of 5×10^{-11} per g residual sensitivity is typical for reproducibly attainable results. Worst axis residual sensitivity can be made smaller with a considerable amount of care. Irregularities of symmetry probably set a practical lower limit around 10^{-11} per g for these resonators.

Perturbation of short term stability is also reduced. The quiescent oscillator shows a two sample, Allan variance, of better than 6×10^{-13} at 1 second averaging time. When the oscillator is perturbed by a sinusoidal 0.5 g input, the stability is 1000 times worse. For the compensated oscillator an improvement by a factor of more than 20 is seen, and is consistent with the compensated static sensitivity.

In the same way, for phase noise sampled at low frequency, a typical improvement is from -70 dB to -98 dB at 15 Hz. Approximate agreement with expected f_m^{-2} is also found.

The results have clear implications for navigation, communications and other systems whose performance depends on low noise high stability frequency sources operating in a non-laboratory environment.

The author wishes to acknowledge the support of the Office of Naval Research under contract N00014-74-C-0061, for portions of this work.

References

1. R. Besson, "A New Piezoelectric Resonator Design" Proc. of the 30th Annual Symposium on Frequency Control, 1976.

2. D.A. Emmons, "A New Rugged Low Noise High Precision Oscillator" Proc. of the 8th Annual Symposium on Frequency Control, 1976.

3. J.M. Przyjemski, "A Compensation Technique for Acceleration - Induced Frequency Changes in Crystal Oscillators" NAECON '78, May 1978; J.M. Przyjemski and P.L. Konop, "Limitations on GPS Receiver Performance Imposed by Crystal-Oscillator g-Sensitivity" NAECON '77, May 1977.

4. The author has recently become aware of the work by Valdois using a mass loaded capacitive accelerometer for frequency shift compensation. Valdois et al., "Crystal Oscillator Compensated Against Frequency Shift due to Acceleration", U.S. Patent 4,100,512, July 11, 1978.

5. J.J. Gagnepain and F.L. Walls, "Quartz Crystal Oscillators with Low Acceleration Sensitivity", NBS IR 77-855, March, 1977.

6. M. Valdois, R. Besson and J. J. Gagnepain, "Influence of Environment Conditions on a Quartz Resonator", Proc. of the 28th Annual Symposium on Frequency Control, 1974.

7. P.C.Y. Lee and Kuang-Ming Wu, "The Influence of Support - Configuration on the Acceleration Sensitivity of Quartz Resonator Plates," Proc. of the 31st Annual Symposium on Frequency Control, 1977.

8. Handbook of Piezoelectric Crystals for Radio Equipment Designers, (P. 105), WADC Technical Report 56-156, October 1956, Wright Air Development Center.

I. HIGH PERFORMANCE CRYSTAL OSCILLATORS IN VIBRATION ENVIRONMENT

   A. DEGRADED STABILITY OF AT CUT RESONATORS

   B. SYMMETRY OF FREQUENCY SHIFT VS. G

II. METHODS OF REDUCING SENSITIVITY

   A. FREQUENCY DETERMINING NETWORK

   B. REQUIREMENTS FOR COMPENSATION

   C. DUAL RESONATOR APPROACH

III. SUMMARY OF RESULTS

FIGURE 1: INTRODUCTION

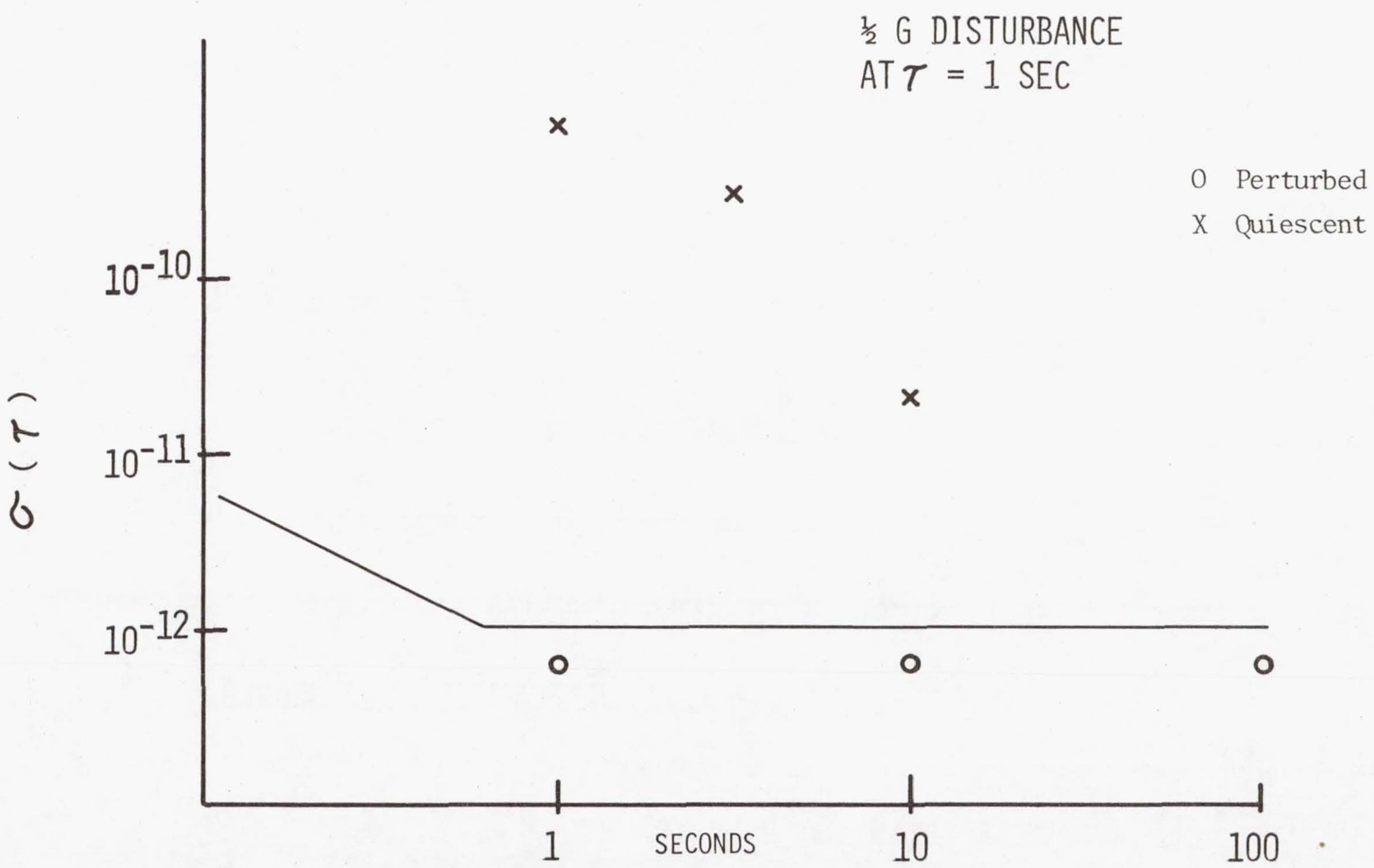
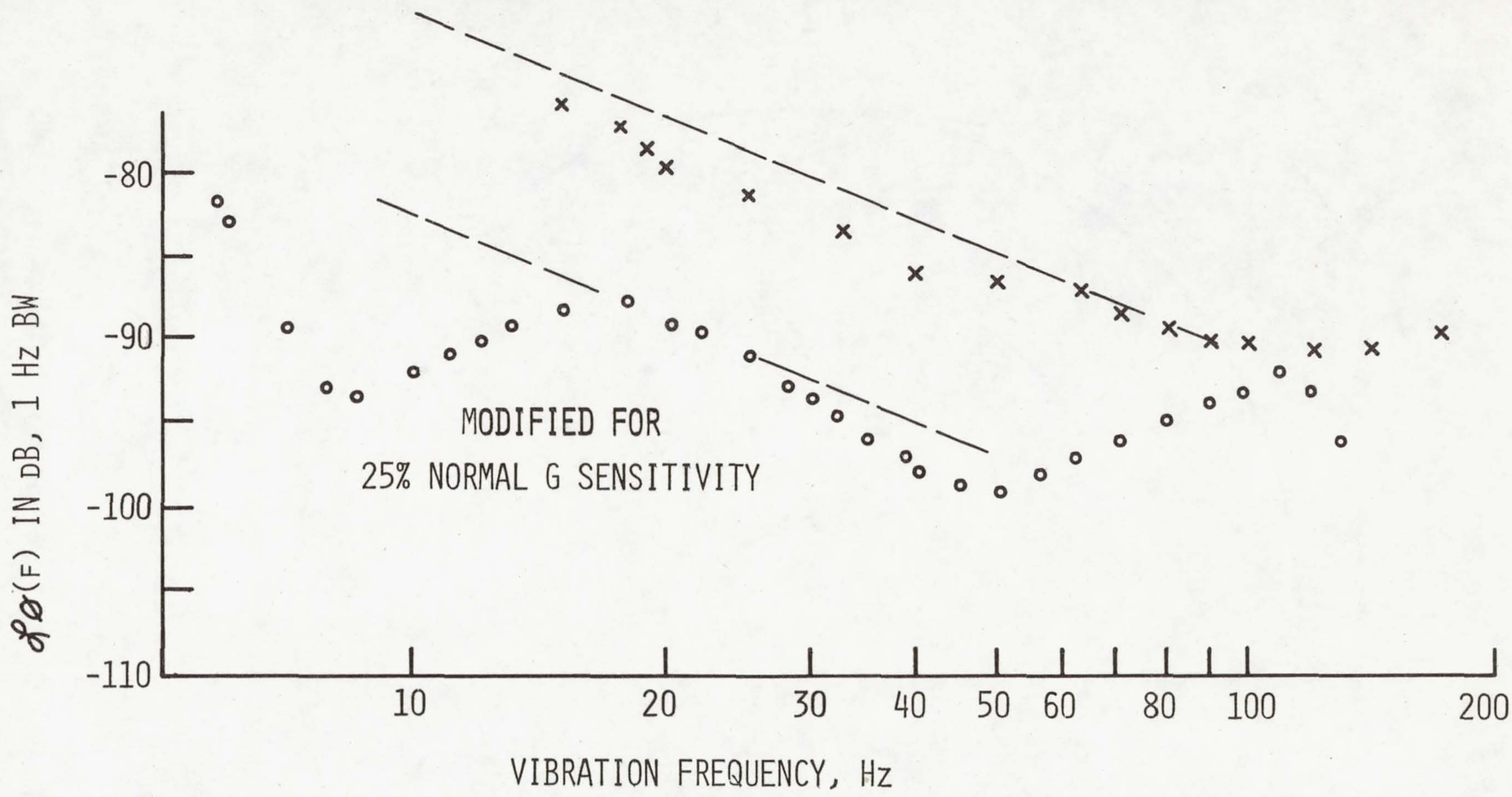


Figure 2. Short Term Frequency Stability



$$\mathcal{L}\phi(f) = (\gamma a f_0 / 2 f_m)^2$$

Figure 3. Side Band Phase Noise, 1 G Sine Sweep Vibration
X. Unmodified MODEL 1000 oscillator, $\gamma = 1.2 \times 10^{-9}$/G
O. Modified for $\gamma = 3 \times 10^{-10}$/G

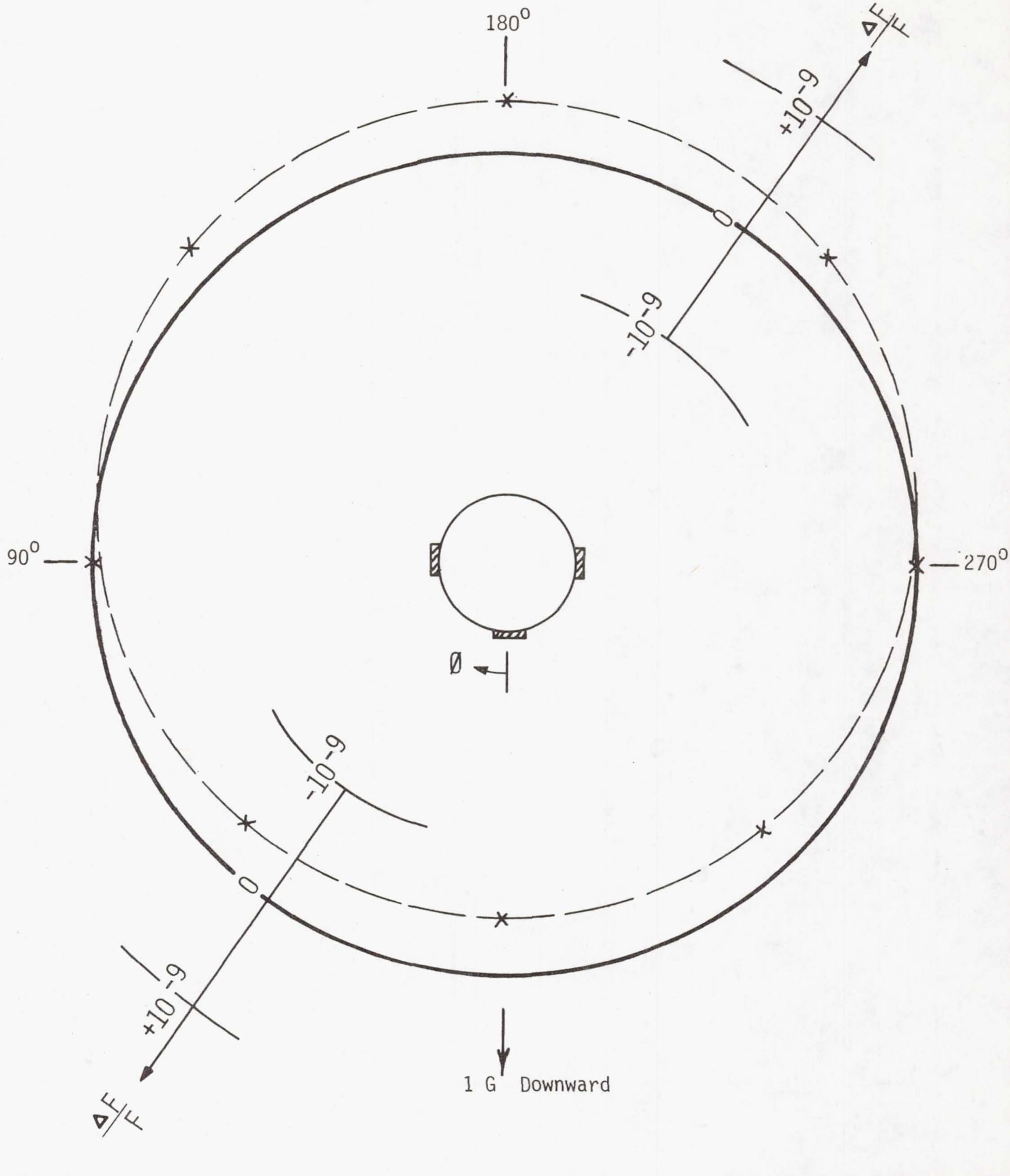


Figure 4. Gravitational Frequency Shift VS. Azimuth for Fifth Overtone 5 MHz Crystal

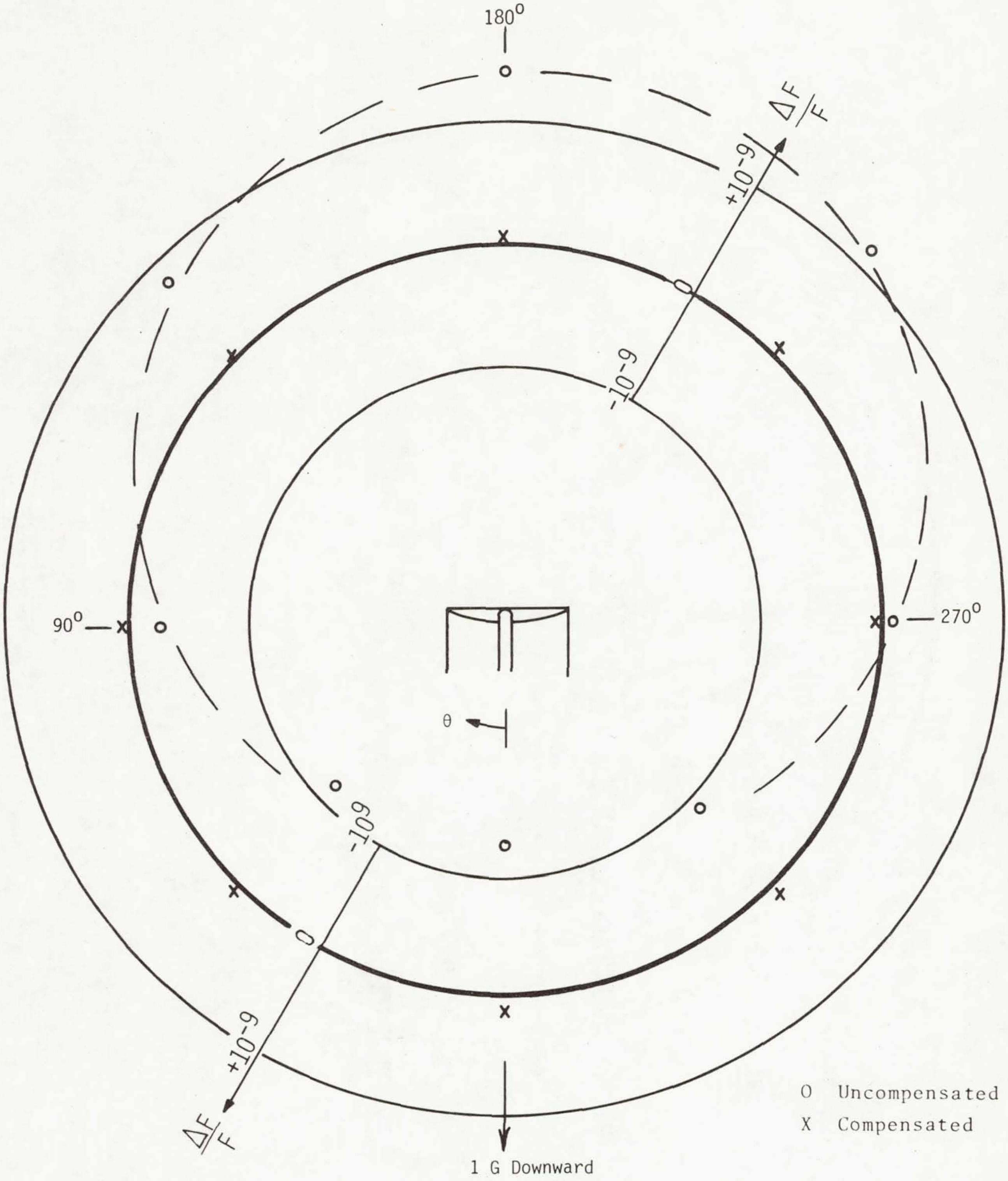


Figure 5. Gravitational Frequency Shift VS. Polar Angle; Compensated and Uncompensated Comparison

REQUIREMENTS FOR IDEAL SYSTEM

1. PERFECT SYMMETRY FOR EXACT CANCELLATION OF VECTOR COMPONENTS
2. LINEARITY WITH G AND WITH FREQUENCY TUNING
3. MODULATION RESPONSE TO LIMIT OF FM BANDWIDTH
4. NO SPURIOUS MECHANICAL PHASE MODULATION
5. NO SPURIOUS CRYSTAL OVEN EFFECTS
6. NO Δ F HYSTERESIS
7. QUIESCENT BEHAVIOR NOT DEGRADED

FIGURE 6: SOME REQUIREMENTS FOR IDEAL COMPENSATION

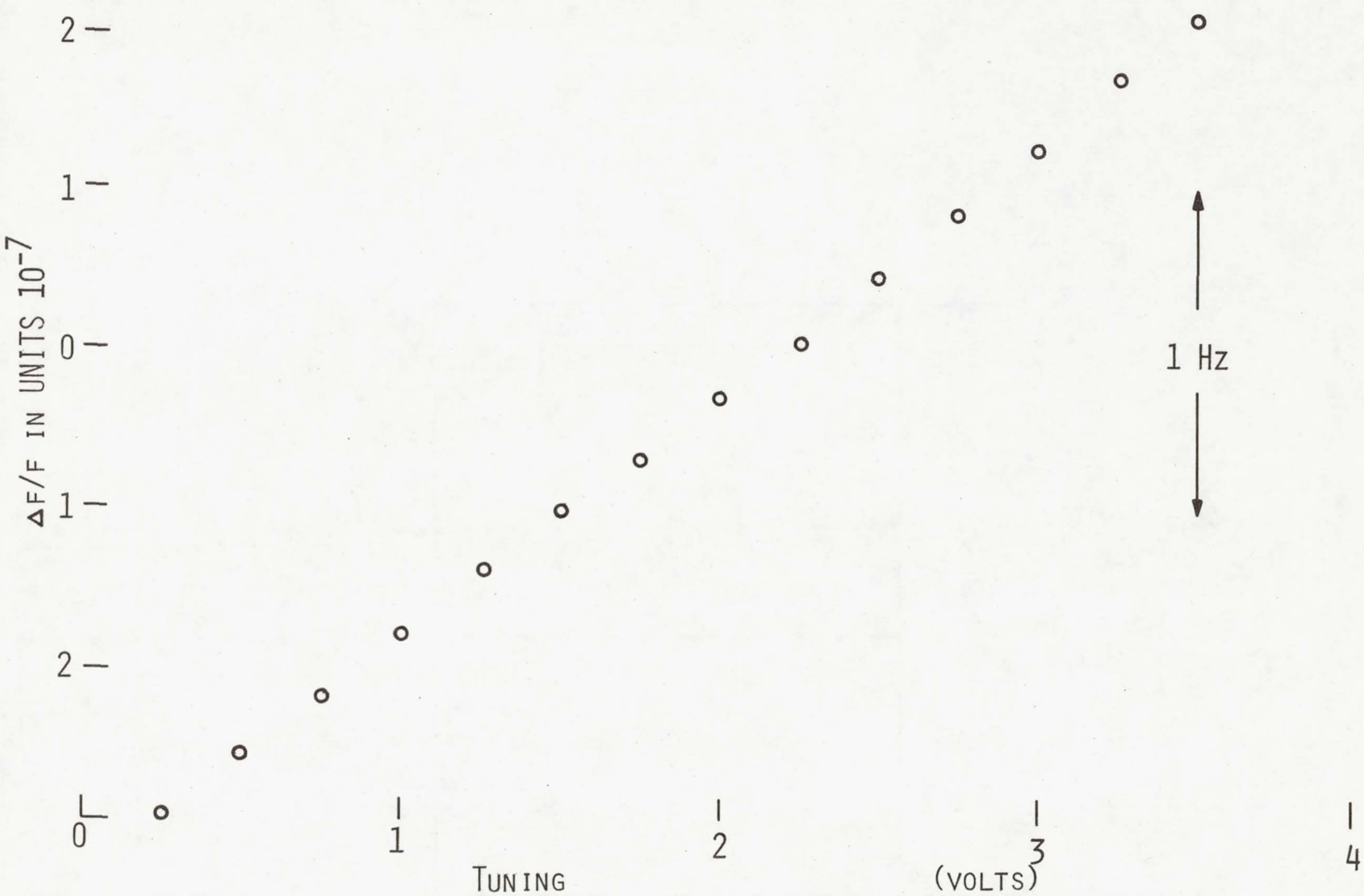
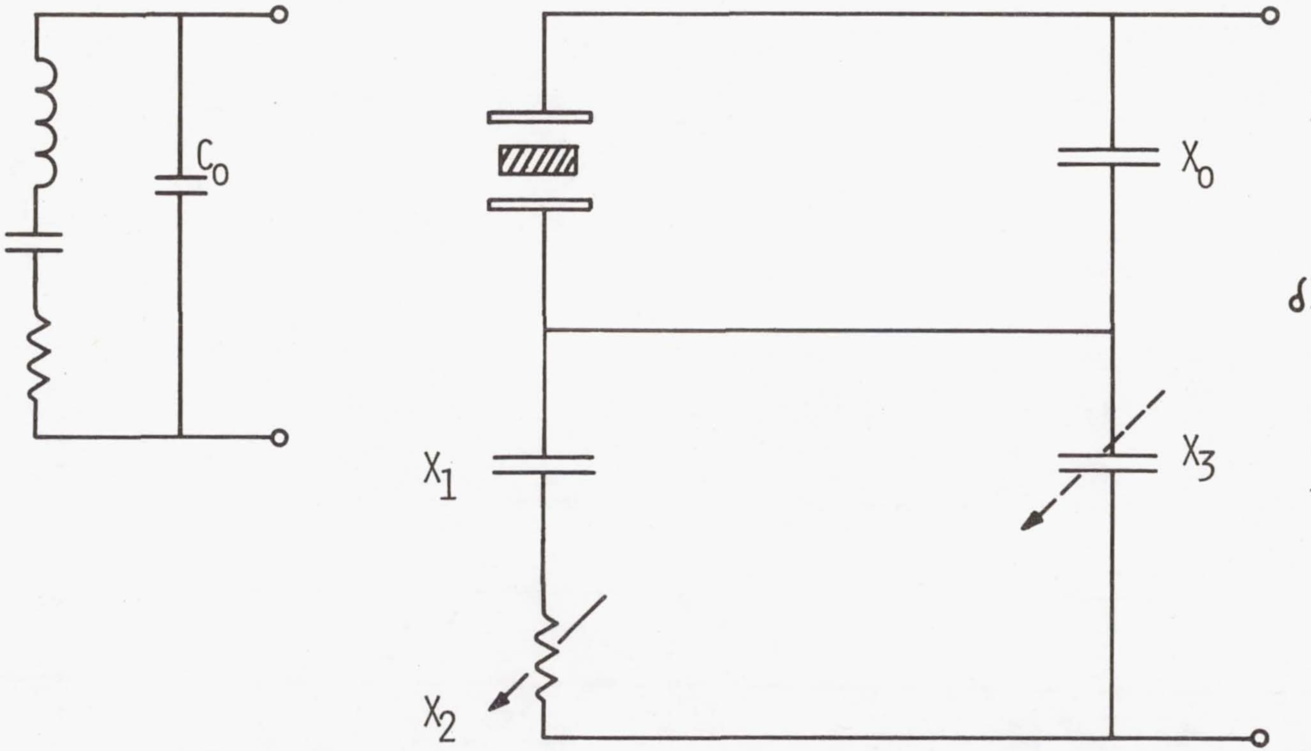


Figure 7. Quartz Oscillator Tuning



$$X_{TOT} \simeq \frac{X_3\,(X_1 + X_2)}{X_3 + X_1 + X_2}$$

$$\delta X_{TOT} \simeq \left(\frac{X_1 + X_2}{X_3}\right)^2 \delta X_3$$

$$\frac{\delta C}{C} = \frac{\delta X_3}{X_3} \propto \frac{\delta s}{s}$$

Figure 8. Frequency Determining Network

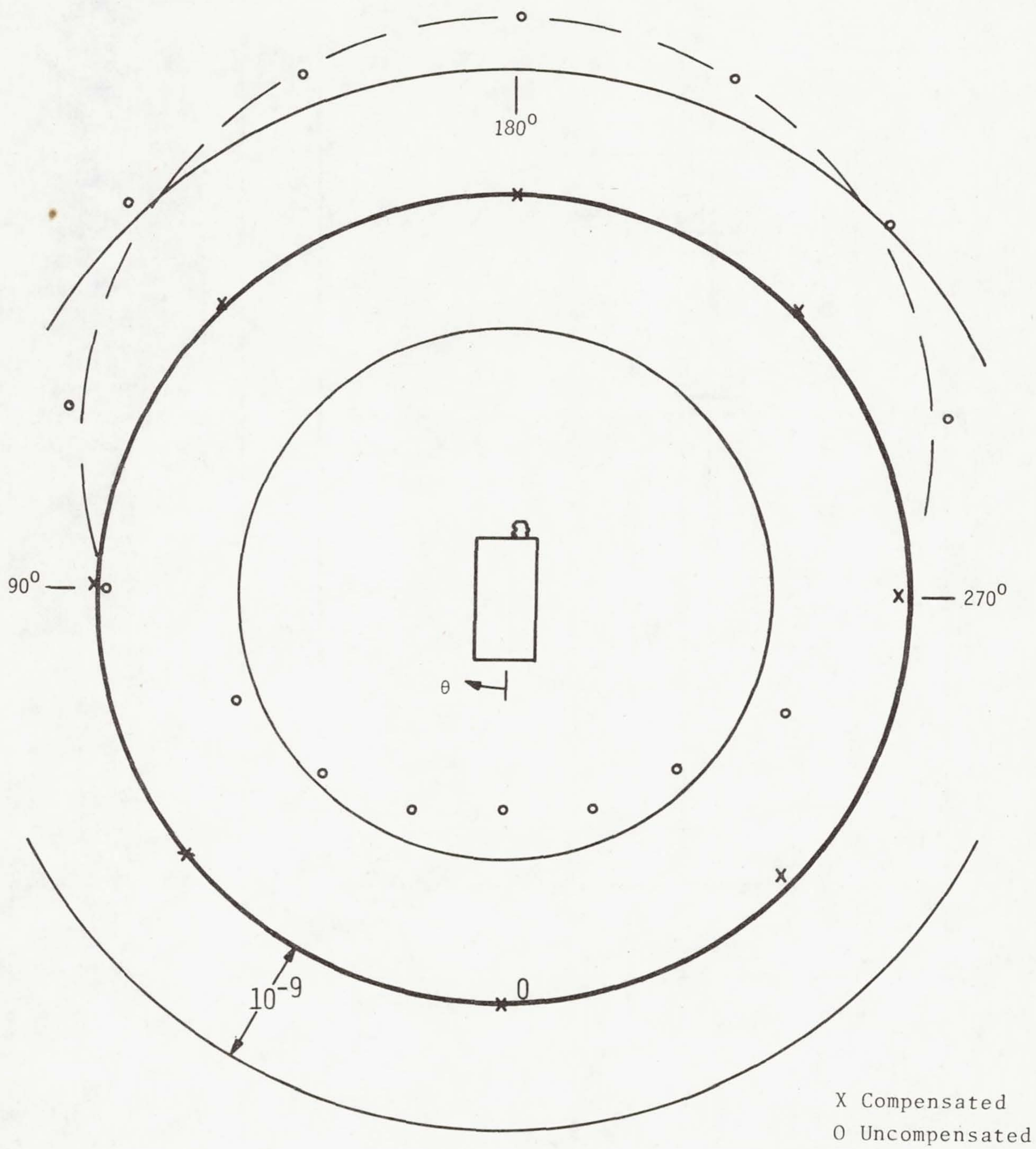
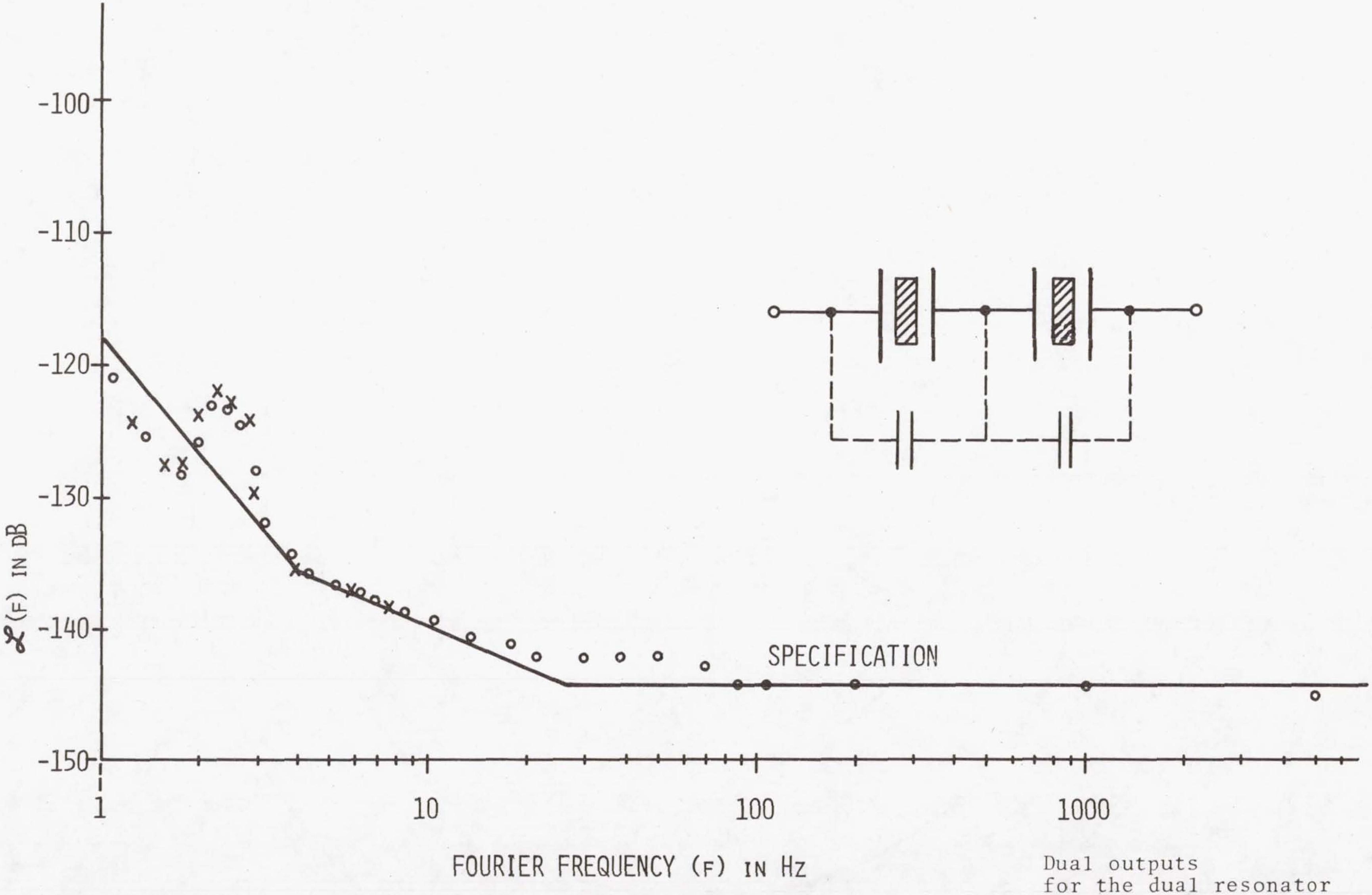


Figure 9. Reduction of G Sensitivity



Dual outputs
for the dual resonator
oscillator

X output A

O output B

Figure 10. Dual Resonator Approach. Phase Noise of Quiescent Oscillator

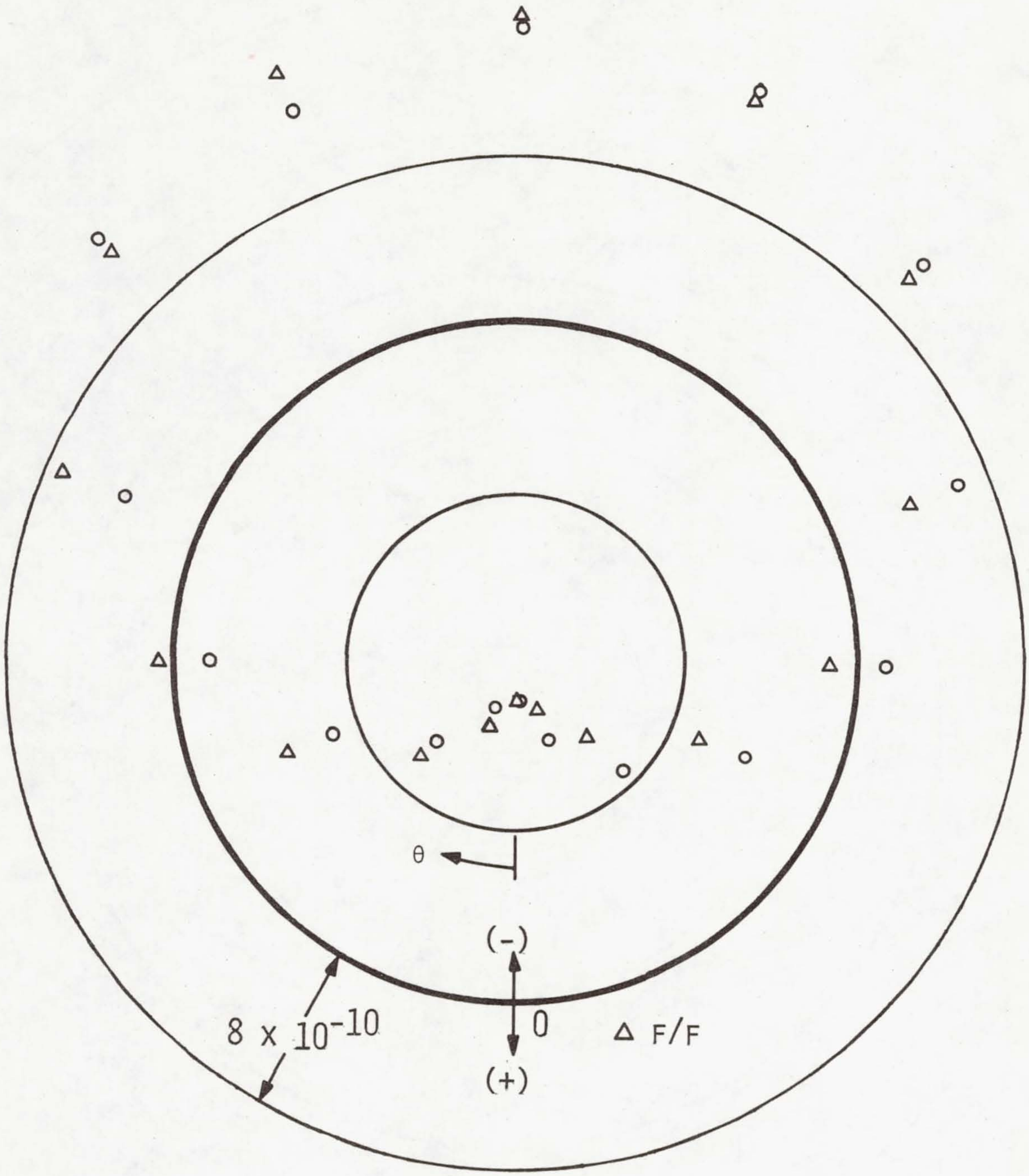


Figure 11. Selection of Matched Resonators

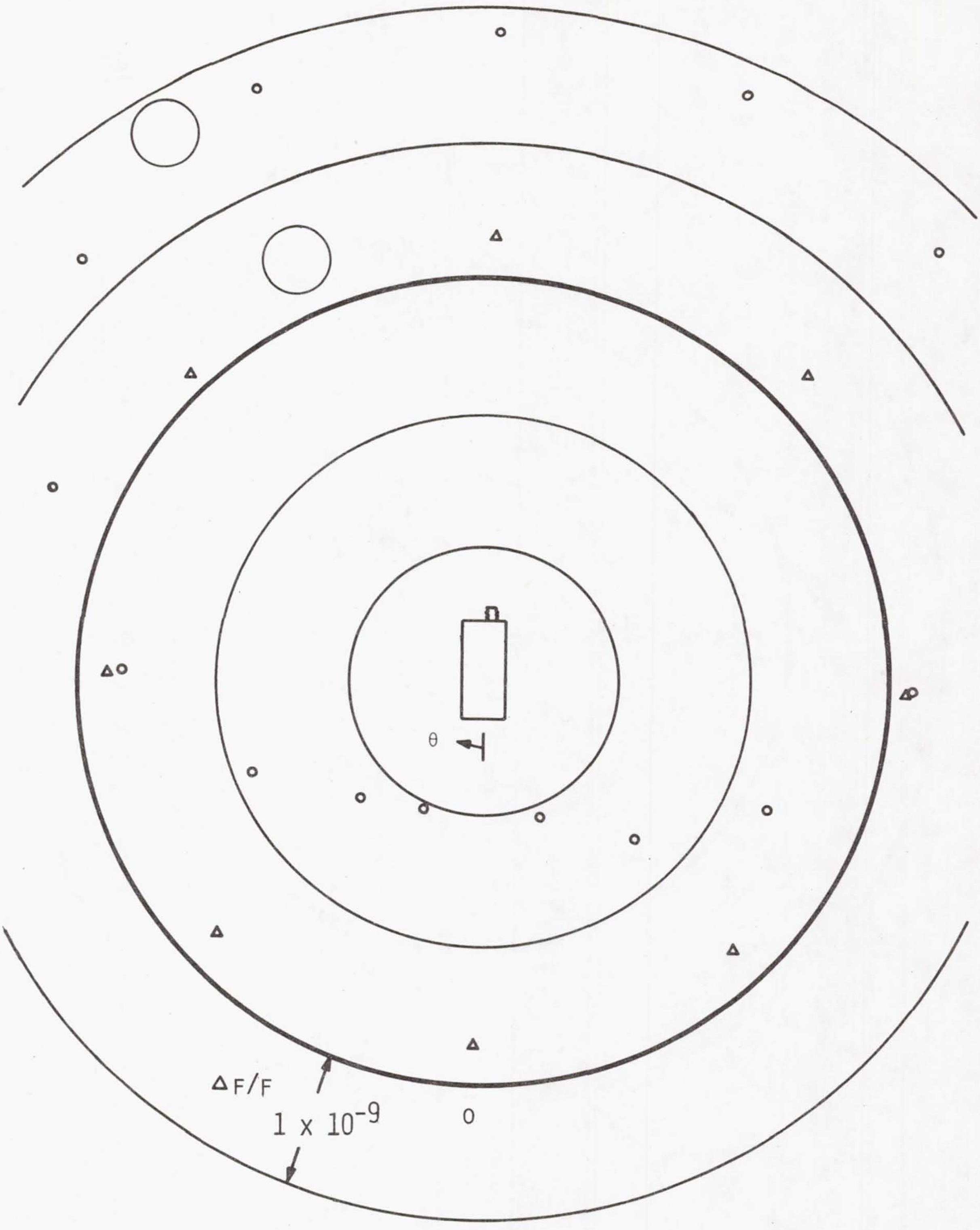


Figure 12. Dual Resonator, 3rd Overtone, Worst Axis Sensitivities
o Single Crystal Response
Δ Dual Crystal Response

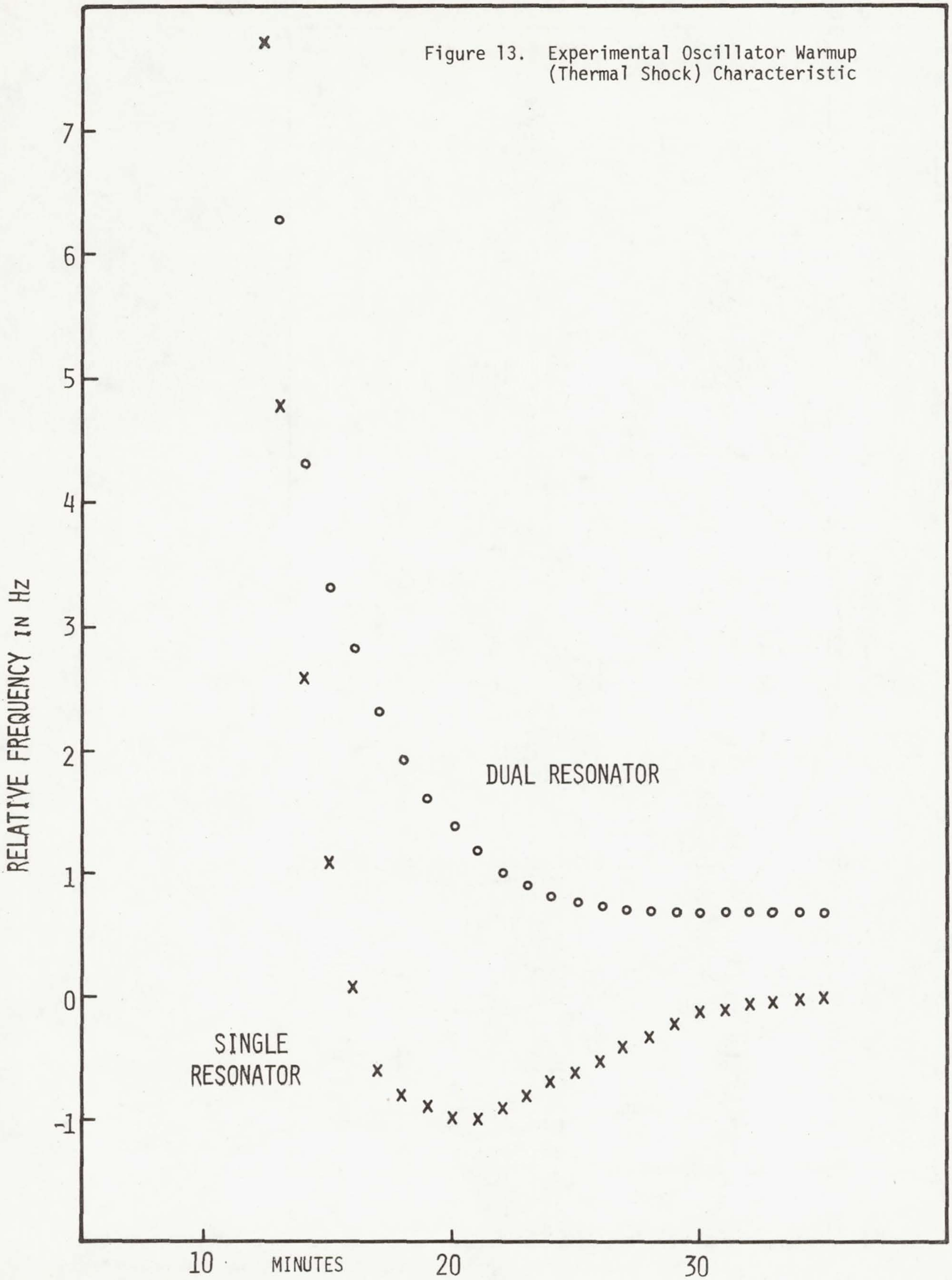


Figure 13. Experimental Oscillator Warmup (Thermal Shock) Characteristic

PARAMETER	UNCOMPENSATED	COMPENSATED
$\Delta F/F$ VERSUS ACCELERATION	$\sim 1.2 \times 10^{-9}/G$	$< 5 \times 10^{-11}/G$
SHORT STABILITY σ (1 SEC) FOR 1/2G INPUT AT τ	70×10^{-11}	3×10^{-11}
PHASE NOISE $\mathcal{L}\,\emptyset$ (15 Hz) WITH 1G SINE INPUT AT 15 Hz	-70 DB	-98 DB

FIGURE 14: SUMMARY. AT CUT 5TH OVERTONE RESONATOR G-SENSITIVITY COMPENSATION

QUESTIONS AND ANSWERS

MR. D. A. EMMONS, Frequency and Time Systems, Inc.:

Could I just make a correction--an error of omission? I think I failed to mention that Presjinski of Draper Labs has very convincingly shown the use of an external accelerometer for g sensitivity reduction. I didn't mean to imply that that was a new technique.

I have also brought with me a desensitized oscillator which is in back beside our portable clock, which fits under the seat of a DC-9.

Page intentionally left blank

PRECISE TERRESTRIAL TIME-A MEANS FOR IMPROVED BALLISTIC MISSILE GUIDANCE ANALYSIS

Eldon E. Ehrsam (Space And Missile Test Center), Stephen A. Cresswell, George R. McKelvey, and Frank L. Matthews (ITT Federal Electric Corporation), Vandenberg AFB, California

ABSTRACT

Significant improvements in ballistic missile guidance performance has necessitated the guidance systems analysis community to adopt more precise timing standards for the evaluation of test data. To date, data timing is achieved at the ground instrumentation sites since the missile dynamics are too demanding for currently available high precision missile borne timing sources. In the past five years as noted in the 1973 PTTI paper, "Precise Timing Correlation in Telemetry Recording and Processing Systems", accuracy improvements in evaluation of guidance systems were expected to exceed the IRIG timing techniques used in time tagging guidance telemetry data. An approach developed by Space and Missile Test Center (SAMTEC) to improve the ground instrumentation time tagging accuracy has been adapted to support the Minuteman ICBM program and has been designated the Timing Insertion Unit (TIU). The TIU technique produces a telemetry data time tagging resolution of one tenth of a microsecond, with a relative intersite accuracy after corrections to better than 0.5 microsecond. Metric tracking position and velocity data (range, azimuth, elevation and range rate) also used in missile guidance system analysis can be correlated to within ten microseconds of the telemetry guidance x, y, z and $\dot{x}$, $\dot{y}$, $\dot{z}$ data. This requires precise timing synchronization between the metric and telemetry instrumentation sites. The timing synchronization can be achieved by using the radar automatic phasing system time correlation methods. Other time correlation techniques such as Television (TV) Line-10 and the Geostationary Operational Environmental Satellites (GOES) terrestial timing receivers are being considered. With the continuing improvement of ballistic missile guidance performance, on-board clocking stability of one part in 10^9 may be required. This increased on-board stability will influence improvements in terrestial time precision and accuracy.

INTRODUCTION

The Air Forces Space and Missile Test Center (SAMTEC), headquartered at Vandenberg Air Force Base, California, manages, operates and maintains the Western Test Range (WTR) and the Eastern Test Range (ETR) in support of Department of Defense (DOD), National Aeronautics and Space Administration (NASA) and other ballistic, space and aeronautical user programs. Figure 1 shows the geographical area of the WTR which extends from the west coast to 90 degrees East longitude in the Indian Ocean and includes instrumentation support provided by the U.S. Navy Pacific Missile Test Center (PMTC), NASA, U.S. Air Force Satellite Control Facility (AFSCF) and the U.S. Army Kwajalein Missile Test Range (KMR). This paper describes the application of precise time synchronization techniques by SAMTEC to provide support for DOD ballistic weapon system user programs with a need to achieve improved ballistic missile inertial guidance system analysis. The SAMTEC has determined through analysis of inertial guidance telemetry data that the missile guidance system clock stability is affected by flight dynamics during the launch and post boost burn. because of this clock stability problem, another means of providing timing synchroniza tion was required. SAMTEC has recently designed and developed equipment and procedures which allow relative time synchronization between telemetry receiving sites to within 0.5 microsecond and the time tagging of recorded telemetry data to a resolution of 0.1 microsecond. Metric radar sites are synchronized to within one microsecond relative between radar sites and telemetry sites. Figure 2 shows the geographical relationship of westcoast uprange (SAMTEC and PMTC) telemetry receive sites and metric radar sites which support ballistic missiles launched from Vandenberg Air Force Base, California.

Background

The need for a more accurate timing capability to time tag telemetry inertial guidance data came as an intrinsic result of the improved guidance performance. Timing accuracy requirements have gradually increased over the past few years from milliseconds in the early 1960's until today when the range user is requesting time resolutions to the tenth of a microsecond. Figure 3 shows the history of timing accuracy improvement implemented at WTR for time tagging of telemetered ballistic missile inertial guidance data.

As recent as only a few years ago, timing provided by the Inter-Range Instrumentation Group (IRIG) "B" codes were considered adequate for post flight reduction of telemetered pulse code modulation (PCM) data. With further improvements in ballistic missile guidance systems, WTR was required to distribute IRIG "A" time code format with a resolution of 100 microseconds. However, even with the higher timing carrier frequencies i.e., IRIG "G" with a 100 kilohertz carrier and a resolution of 10 microseconds, timing accuracy and precision were still found to be inadequate for the demands of inertial guidance accuracy analysis. This is primarily due to the methods of recording the timing on independent tracks of an analog tape recorder in parallel with the predetection recorded PCM telemetry data. Tape recorder electronics [1] introduced phase shifts in the recorded IRIG time code carriers. Physical placement (alignment) of the reproduce heads on the tape recorders would also introduce variable time biases from machine to machine. Difficulties in predicting system time delays, not only in the telemetry data but in the IRIG time distribution equipment resulted in the search for a more accurate and reliable means of correlating the same telemetry time tagged data recorded at multiple sites.

The initial investigation by the SAMTEC resulted in the design and implementation of a Time Correlation Generator (TCG) system. The TCG uses a one pulse per second (pps) strobe from the telemetry site local time code generator and then inserts a recognizable pattern into the postdetection recorded telemetry bit stream. Through post launch analysis of this data, the inserted pattern is correlated with the IRIG millisecond markers and a relative timing bias correction factor is calculated and applied to the inertial guidance data during post launch processing. Through the use of the TCG system SAMTEC analysts were able to measure inertial guidance internal timing instabilities which were not previously measured during missile flight. However, the TCG system had two basic limitations, the resolution after data processing was limited to plus or minus one bit time (about 3 microseconds for the prime missile system user) and the repetitive frequency of time correlation (once per second) introduced aliasing.

The ballistic missile user has specified new timing requirements to time tag all telemetry data referenced to the time it was transmitted from the missile. This requires the synchronization of all uprange telemetry sites to within 3 microseconds relative and all metric radar sites to within 10 microseconds relative of each other and the telemetry sites. This has resulted in the latest uprange WTR configuration for time synchronization and distribution. The user's timing requirement was satisfied through

two actions by SAMTEC. First, time tagging of the telemetry data and time synchronization between telemetry sites was accomplished through the design and implementation of a new SAMTEC development designated the Timing Insertion Unit (TIU) and secondly, the metric radar time correlation will be accomplished through the use of intersite time correlation technique using the radar Automatic Phasing System (APS).

Timing Insertion Unit

The time synchronization accuracy requirements necessitated the development of a new operational approach to the problem of accurate time tagging of missile events. These events were to be tagged with respect to the time of transmission rather than time of data receipt. This requires not only synchronization of time between telemetry sites but also an accurate knowledge of the missile location during flight. This latter requirement is solved through the use of inertial guidance position and velocity data in terms of x, y, z and $\dot{x}$, $\dot{y}$, $\dot{z}$, contained within the telemetered PCM data stream. But additional confirmation of the missile position is required for guidance analysis. Therefore, accurate metric radar data, correlatable to the telemetered PCM data, is also required. From these factors, timing support requirements were fixed to within 3 microseconds between telemetry sites and to within 10 microseconds between radar and telemetry sites. Resolution of the telemetered PCM data and inserted timing was fixed at one tenth of a microsecond.

The goals of the design require that all known sources of error such as telemetry propagation delay, receiving system delay, etc. be minimized. In addition this should provide the necessary information to accomplish a relative time correlation between telemetry sites. The TIU units, installed at each telemetry site, derive their time stability from cesium frequency standards traceable to the U.S. Naval Observatory (USNO) Coordinated Universal Time (UTC). The traceability to USNO/UTC is established through a SAMTEC Precise Time and Time Interval (PTTI) calibration program.

The basic operational concept of the approved TIU design recognizes the telemetered frame sync pattern of the PCM frame or subframe. After frame or subframe recognition it then waits for the leading edge of the first PCM bit following frame sync to occur and strobes binary coded decimal (BCD) time (in hours, minutes, seconds, milliseconds, microseconds and hundreds of nanoseconds) into a storage register. The time word is then strobed from the storage register into the PCM data stream at a known and predetermined data word location. The PCM/TIU data is then outputted in a serial format for recording and subsequent data processing.

A functional block diagram of the TIU application is shown in Figure 4. The TIU interface consists of a PCM bit synchronizer which reshapes the raw PCM telemetry data and generates a zero degree clock, a frame decommutator which provides frame sync recognition input, a time code generator for providing BCD timing input, a cesium frequency standard for providing a stable five (5) megahertz (MHz) input frequency to the time code generator, and associated power supplies and output driver circuits.

Since all inserted BCD timing from each TIU is referenced to the arrival of the same telemetry data bit, and the position and velocity information in the terms of x, y, z, and $\dot{x}$, $\dot{y}$, $\dot{z}$ of the transmitting ballistic missile is well known; it becomes a routine processing computation to first, correct the inserted time tag to the time of data transmission and secondly, to determine relative time from one telemetry site to another. If required, correction of the inserted time at any site relative to any other site can be computed. Thus, through the application of the TIU and post launch processing of data preoperational telemetry site synchronization of time is not required. It is only necessary to have each supporting telemetry site provide overlapping telemetry coverage in order to effect an intersite time correlation of the post launch processed telemetry data.

Figures 5 and 6 show typical telemetered transit time corrected TIU measurements of on-board inertial guidance system clock instability with a peak-to-peak jitter of 40 microseconds. The two sets of data were recorded over the same flight time interval from two independent WTR telemetry acquisition sites with overlapping coverage. The precision and accuracy of the TIU is one part in 10^7.

Metric radar timing correlation is accomplishd using the automatic phasing system at uprange locations.

Radar Automatic Phasing System

The WTR metric radar systems utilize the APS to prevent "beacon stealing" or multiple radar returns, when two or more radars are tracking a single target. The radar phasing system designates the radar under test into a transmit time slot and maintains the radar transmitter in that time slot regardless of its target range; this prevents an overlapping of the radar returned signals. Figure 7 is a simplified block diagram of the APS used at the WTR. As shown, the radar pulse repetition frequency (PRF) of 160.0864 pulses per second is derived by division of a 5.24571328 MHz signal which is synthesized from a 5 MHz cesium frequency standard signal. The time of occurrence (TOC) of the first pulse of the radar 160.0864 PRF is controlled by a time code generator/ synchronizer

also driven by the cesium 5 MHz frequency reference. Thus, if all radar site cesium frequency standards are "on time" and all radars are assigned to the same transmit time slot, the transmit times of all radars will occur at the same time epoch.

If relative time correlation between two radar sites cesium frequency standards is required as illustrated in Figure 8, it is only necessary that each radar system range on the others transmit pulse. The relative range measurement difference now represents the relative time difference between the two site cesiums frequency standards. To synchronize the two sites it is necessary to adjust one of the cesium frequency standards until both radar systems measure the same range. Time correlation using this method has a resolution accurate to within the measurement accuracy of the radar which is less than 50 feet, or approximately 50 nanoseconds in time, and assumes only that the two-way radar range transmission times are the same and that internal system time delays within the radars are known. Typical results of APS measurement are within plus or minus one microsecond as compared against a portable traveling clock. This clock also provides the traceability to USNO/UTC.

The method of time synchronization between radar and telemetry sites at the WTR has been greatly simplified since one of the cesium frequency standards located at the Pillar Point Air Force Station, California provides a common frequency reference to colocated telemetry and metric radar systems.

Backup Time Correlation Considerations

The use of the metric radar APS to achieve the 10 microsecond relative time correlation between metric radar sites and telemetry sites has been demonstrated. The APS availability for time correlation measurements cannot always be assured because of the higher priority operational system commitments. To circumvent this, SAMTEC is evaluating the use of other backup time correlation systems i.e., TV Line-10, Geostationary Operational Environmental Satellite (GOES), and other available systems.

The TV Line-10 relative time correlation application is jointly being evaluated between Camp Roberts U.S. Army Satellite Communications Station and the SAMTEC Precision Measurement Electronics Laboratory (PMEL) Vandenberg Air Force Base, California. Preliminary data shows better than one microsecond resolution when compared with the PMEL cesium frequency standard and LORAN-C. In addition, SAMTEC has obtained two GOES satellite timing receivers for evaluation.

A secondary benefit of either system is that it will provide SAMTEC performance analysts a means to trace the cesium frequency standards instability (drift) with USNO reference clocks. This capability provides a more effective method to evaluate cesium frequency standard performance and calibration intervals.

Conclusions

SAMTEC has demonstrated that through the use of the Timing Insertion Units and Automatic Phasing Systems, that relative intersite timing can be correlated and controlled to less than one microsecond for telemetry and metric radar sites. Although further improvements to the WTR terrestial timing system are possible, it appears that until the on-board inertial guidance clock stability is improved to one part in 10^9, the WTR terrestial timing system will not be a limiting factor.

REFERENCE

1 Pickett, R. B., Matthews, F. L., "Precise Timing Correlation in Telemetry Recording and Processing Systems" Proc. 5th Precise Time and Time Interval Planning Meeting Goddard Space Flight Center, Greenbelt, Md., December 1973. NASA Publication X814-74-225.

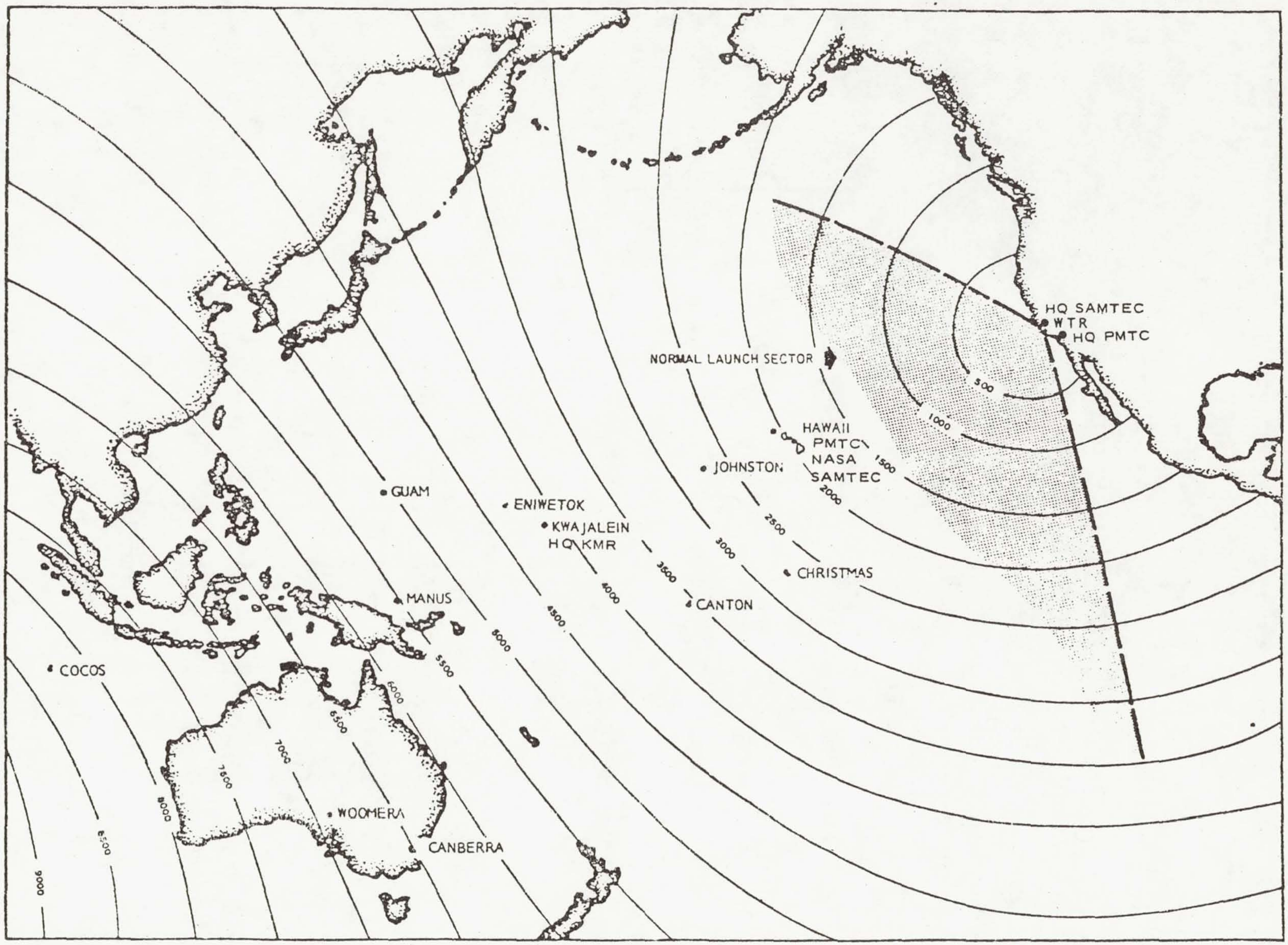


Fig 1 Western Test Range Sector

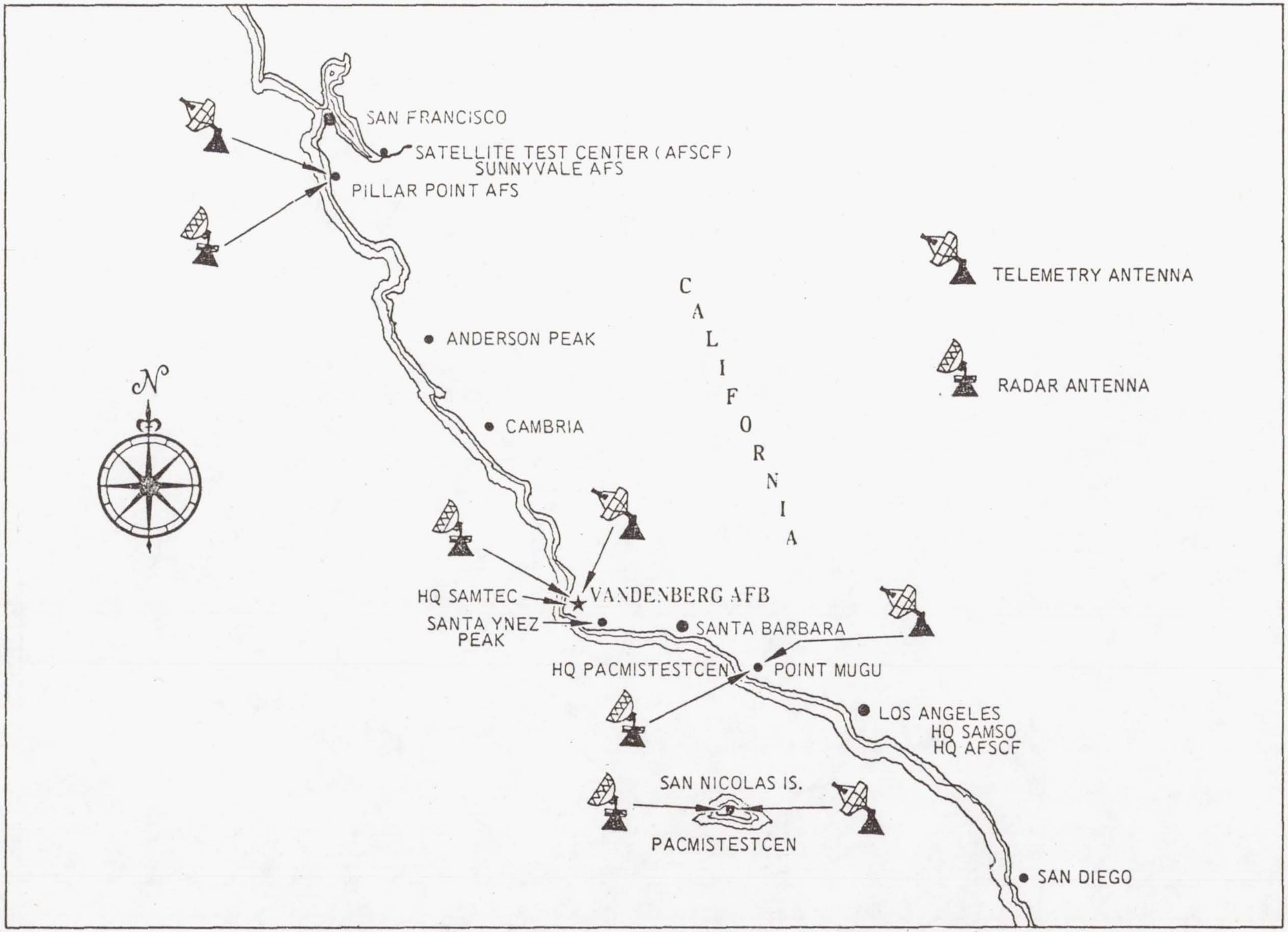


Fig 2 Western Test Range Uprange Facilities

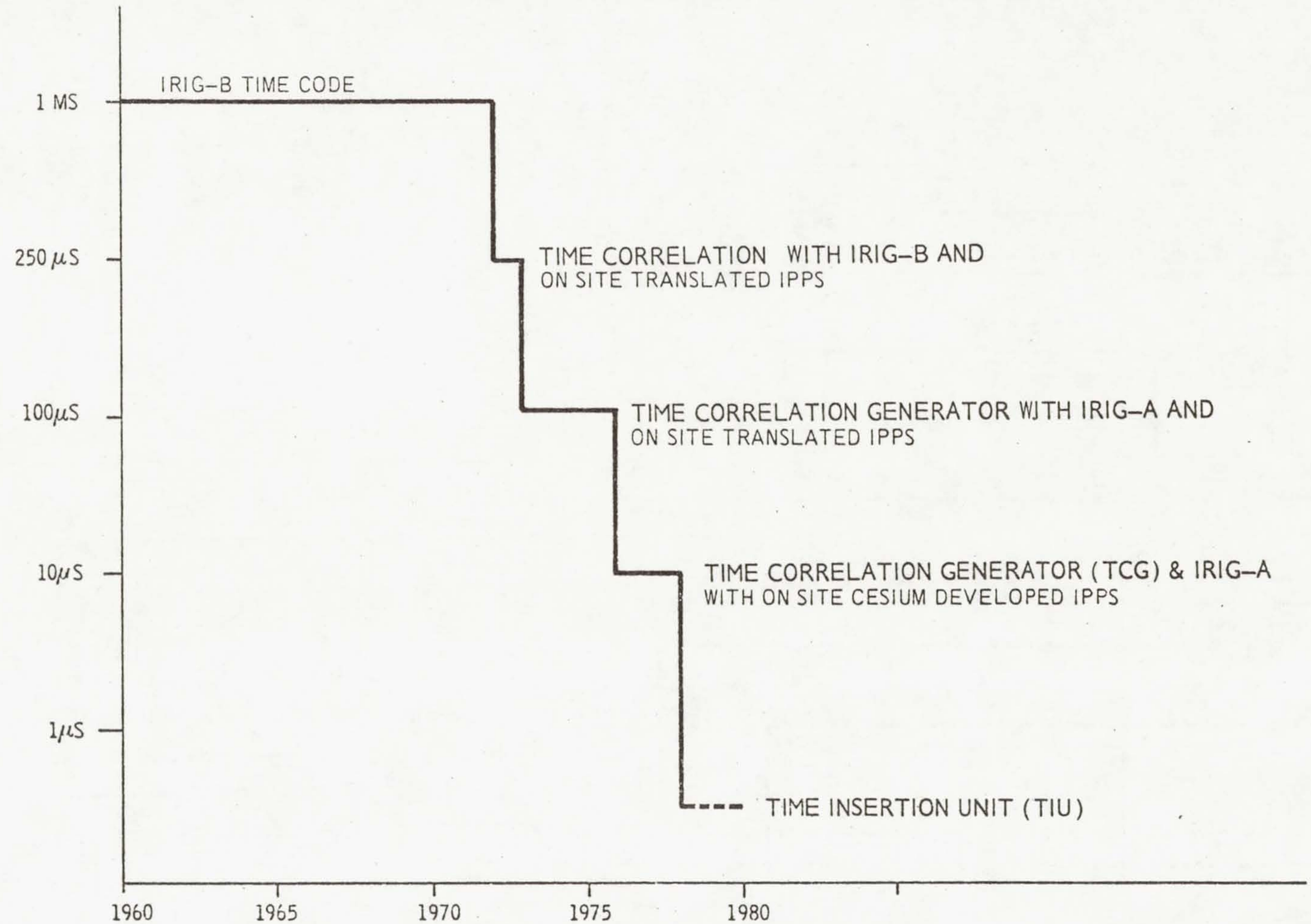


Fig 3 Histogram of Telemetry Data Time Accuracy Improvement

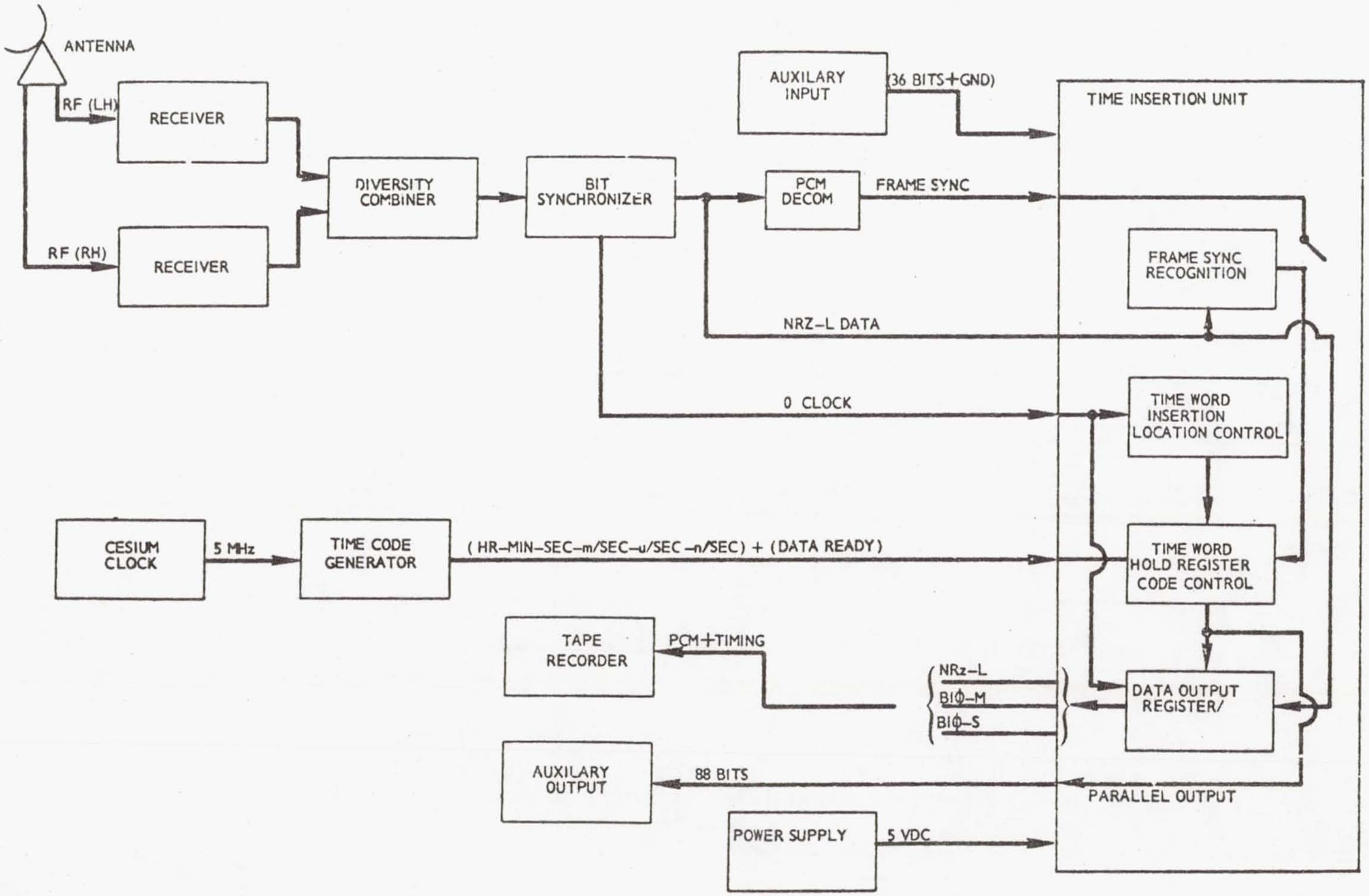


Fig 4 Functional Block Diagram of Time Insertion Unit Application

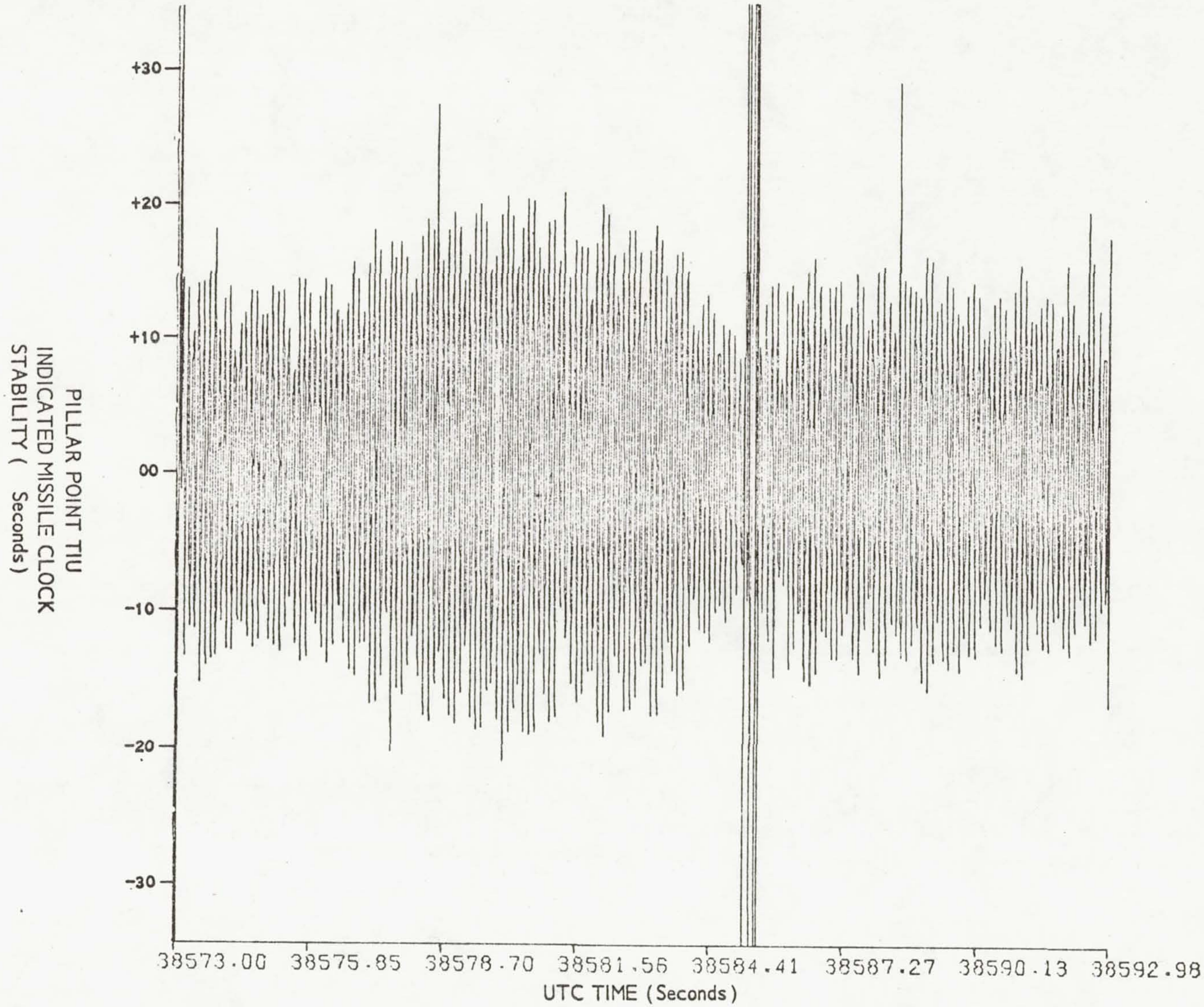


FIG 5 PILLAR POINT*

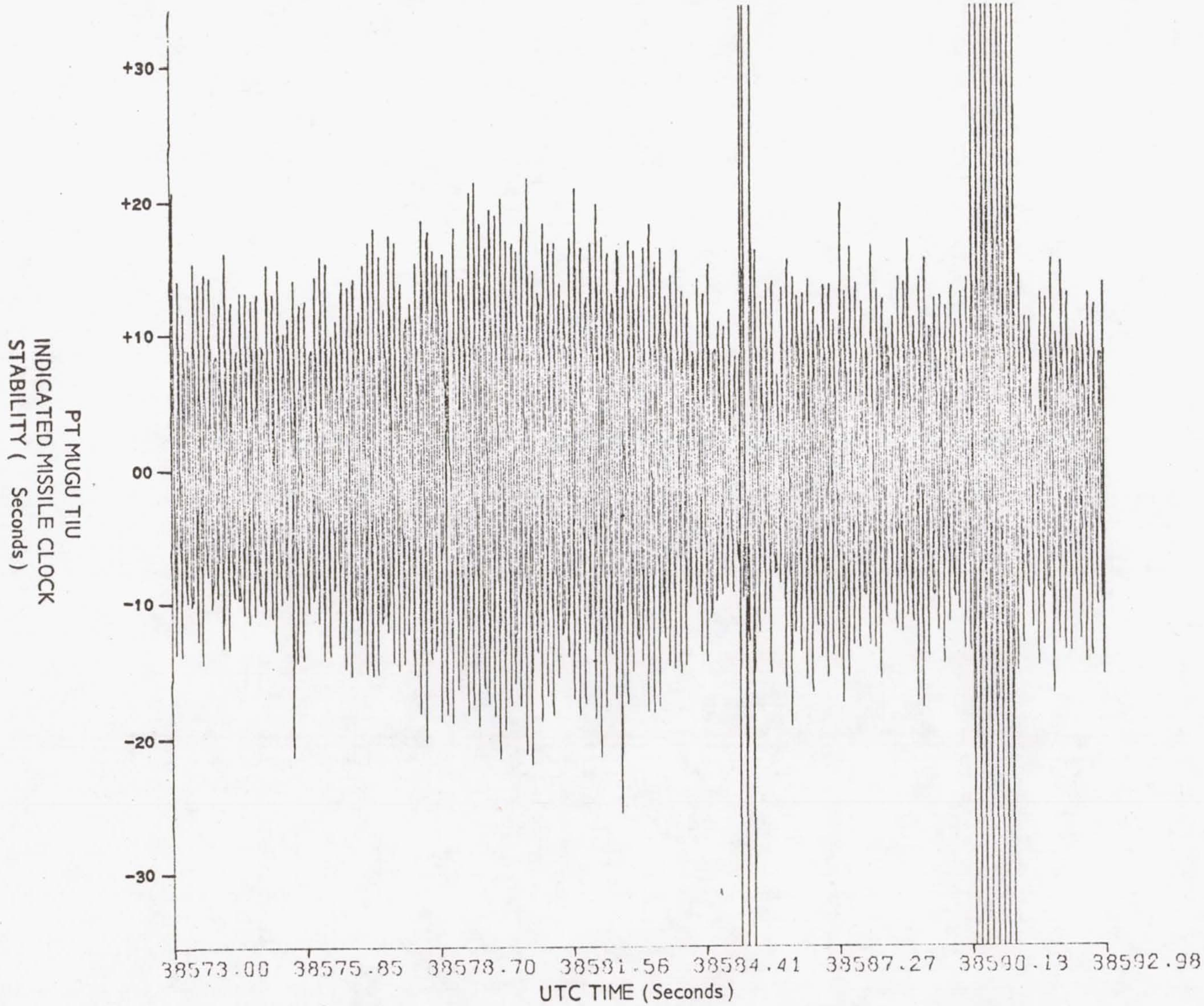


FIG 6 POINT MUGU*

* Includes acquisition site hardware anomalies.

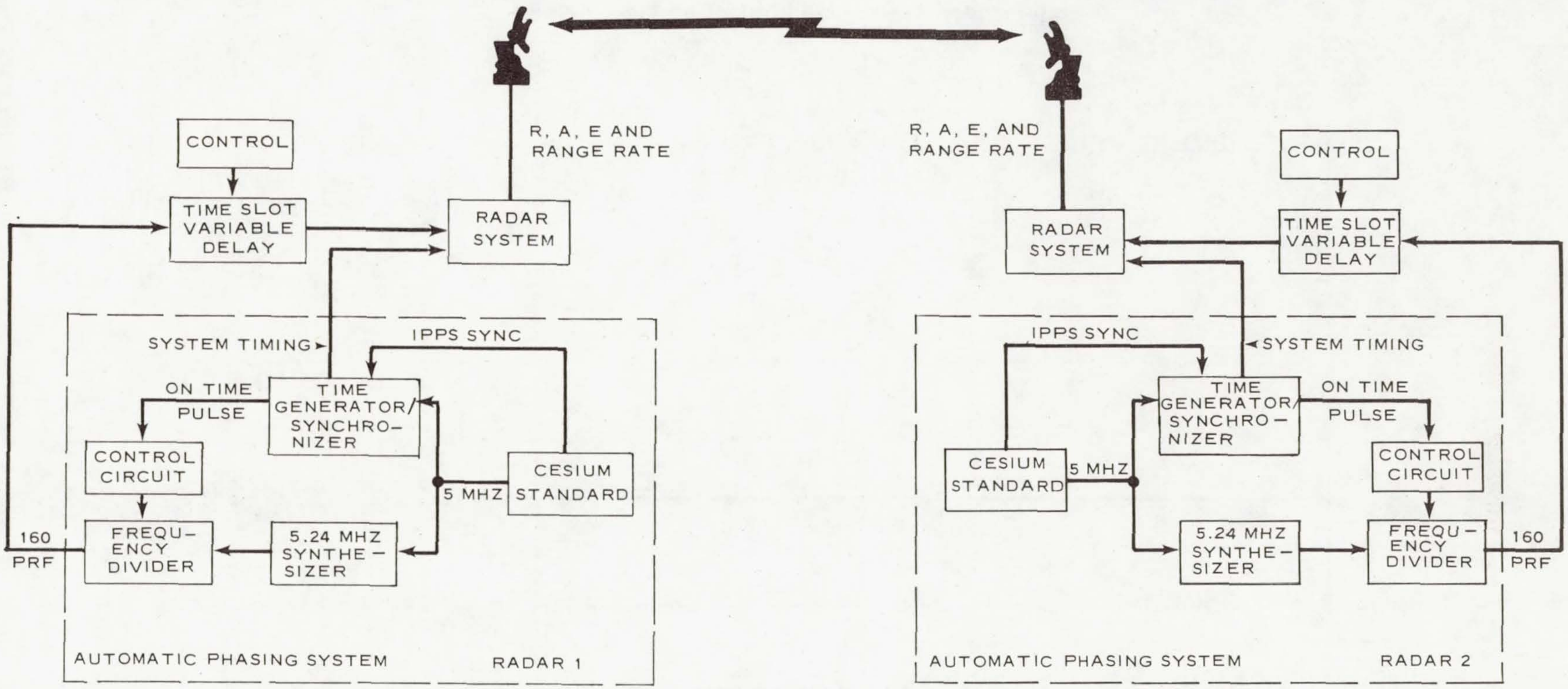


Fig 7 Operational APS block diagram of typical time correlation measurement.

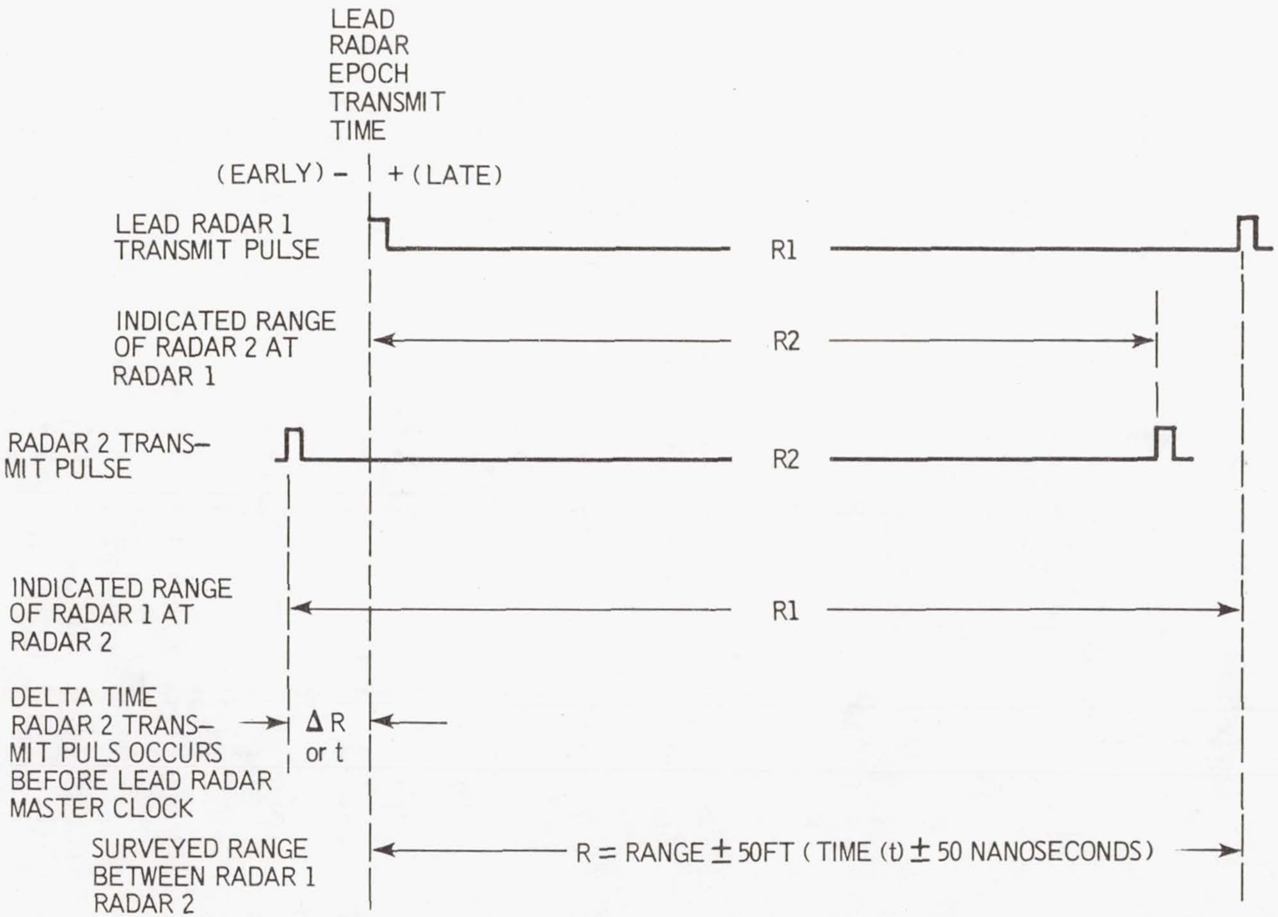


Fig 8 Time Correlation between two radar sites.

QUARTZ CRYSTAL AND SUPERCONDUCTIVE RESONATORS AND OSCILLATORS

Raymond J. Besson

Ecole Nationale Supérieure de Chronométrie et de Micromécanique
La Bouloie - Route de GRAY -
25030 BESANÇON Cédex - FRANCE -

ABSTRACT

Recently, tremendous advance has been made in the field of "quartz crystal and superconductive resonators and oscillators". A review of general related concepts in the field is presented whereas recent achievements are discussed indicating possible trends for immediate future.

INTRODUCTION

This paper, dealing with a very wide subject, adresses various aspects which are not always considered as connected. Our purpose will rather be to provide general concepts and attempt to dig out some underlying ideas or to indicate some tendancies that could be of interest in the near future. Under those conditions details will often be ignored ; we will not pretend to exhaustiveness but rather give a particular point of view, referring when necessary to original or more complete papers. In recent years tremendous advances have been made in oscillators conception (1),(2), in crystal units design and achievement (introduction

of new crystal cuts, new resonators and transducers (3),(4),(5),(6),(7))and in superconducting oscillators and resonators (8),(9),(10) (achievement of noise floor of $3\,10^{-16}$ for times longer than 10 s). Quartz crystals and cavities will be adressed the same way (quartz resonators are ultrasonics cavities). Those advances have been obtained through fundamental and technical achievements as well, and both were needed <u>at the same time and in every related field</u> (resonator theory, material investigation, non linear phenomena, resonator design, crystal orientation, surface decontamination, encapsulation processes and so on). Each aspect has to be considered as important for desired performances (including environmental features).

Piezoelectric resonators and superconducting cavities will first be introduced together with some new concepts in resonator design. This will demonstrate that some interconnection between various fields is needed to really obtain improvements. Some recent results will then be given and a conclusion will be discussed to indicate which improvements can reasonnably be expected in the near future.

Next section will give a general presentation of piezoelectric resonators and their evolution pointing out the main difficulties successively met and solved as the technique evolves toward satisfying requirements.

I - Importance of the resonator's part

In general a resonator is some device possibly vibrating at various frequencies and exhibiting various losses which makes the vibration look like a damped oscillation versus time. A negative dynamic resistor can, for instance, compensate those losses but a general rule for success is to start with the minimum losses, i.e.the best possible Q factor, and perturb the resonator as little as possible. Electromechanical coupling makes piezoelectric resonators very attractive since it provides easy excitation and detection of some resonances

(nevertheless, as a first approximation, piezoelectricity does not drastically change the problem of the vibrating body). The extremely high Q factors achieved with superconducting cavities make them very attractive resonators which have proved to allow the best stabilities. The possibility of using adequate piezoelectric resonators at low temperatures can be very promising since the Q factor is expected to increase largely.

1/ Brief presentation of piezoelectric resonators

Piezoelectric resonators are solids of given configuration, shape and dimensions prepared from piezoelectric crystals under precise control of geometry and orientation. Only certain vibrations are piezoelectrically driven. Electrodes provide the electric field necessary to excite the desired mechanical resonances (and other through non linear effects). Though, basically, the same phenomena are involved a great difference exists between bulk devices and S.A.W. devices. S.A.W. devices will not be considered in details here, nevertheless it is interesting to point out their interest for high frequencies and probably for short term stability since the energy applied on a S.A.W. resonator can be higher before non linear effects occur. Because of its many advantages, quartz material has not been seriously challenged in most applications. However lithium tantalate and lithium niobate have proved to be attractive in certain applications (filters, S.A.W. devices ...).

Since an excellent review paper on quartz resonators is available (11) and provides all desired references up to 1974, in the present review we shall call attention to some important aspects emphasizing the most evalutionary ones. From the usual equivalent circuit valid around a resonant frequency, let us simply retain here that the intrinsic Q_i factor of a crystal is given by:

$$Q_i = \frac{\bar{c}}{\omega \eta_S}$$

where - $\bar{c}$ is the rotated elastic coefficient (piezoelectric effect included)

- η_S is the viscosity constant.

Then the Q of an actual resonator depends :

- on the intrinsic Q_i i.e. on the actual cut, frequency and η_S (a BT cut has basically 1.6 times the Q factor of an AT cut)

- on the construction of the resonator (ideally one would aim to obtain one resonant frequency only and introduce no other losses than material losses).

2/ Frequency temperature behaviour. Various cuts

For a bulk resonator, high frequency, thickness shear vibrating the frequency f at a temperature T between -200°C and +200°C is given by :

$$\frac{f - f_o}{f_o} = a\,(T - T_o) + b(T - T_o)^2 + C(T - T_o{}^3$$

where :

$f = f_o$ for $T = T_o$

a, b, c are functions of the rotated material constants <u>for a given mode of vibration</u>.

The singly rotated AT cut (C mode vibrating) is one orientation for which : $a = b = 0$

This cut will provide us with an interesting behaviour versus temperature since two "turn over" points are available on the plot of frequency versus temperature thus giving rise to temperature stabilization. Let'us assume now that additionnal interesting properties are needed. For instance minima non linear effects (13),(14) or indepen-

dance versus stresses in the plane of the cut (4) are desired. One more degree of freedom must be used thus leading to doubly rotated cuts which have already a long story since the first patent was awarded in 1936. Nevertheless the most important advances are recent (and largely due to the efforts of A. Ballato, E. Eernisse and others) and an important industrial development has already been done (the investigators in industry will forgive me not to cite them all).
Similar considerations to those developed for bulk resonators can be developed for S.A.W. devices (15) and give very promising cuts as a result.

3/ Investigating the modes of motion

If the crystal vibrates, the corresponding waves propagate and may form standing wave patterns for given frequencies. Those frequencies are resonant frequencies. The corresponding patterns are called resonance of motion. Complete understanding and investigation of those modes are needed for any improvement in quartz crystal design. Usually, the radiofrequency spectrum of a resonator is very complicated (but still does not comprise vibrations which do not give rise to an electric resonance (16)),and it's complete investigation needs theoretical developments(17)(18)(19)together with accurate analysis by X ray topography scanning electron microscopy, interferometry, holography or so. In addition linear considerations are not sufficient at all, since linear effect would not account for coupling between modes (16),(19), amplitude frequency effects (14),(20), and so on.

4/ Evolution of the resonator design

Piezoelectric resonator has a long story since it began in 1918. Between the two world wars resonators have been developed at a low industrial rate, rather on an artistic basis. World war II has been

the great booster of quartz industry. The most important improvements have been performed by R.A. Sykes who introduced the universal use of coated units in the 1948 and A.W. Warner who achieved in 1952 the design which is still used now without any major change. Then, for a time, improvements have appeared as asymptotical and interest has decreased. Nevertheless, since quartz oscillators are present in almost any frequency control equipment and are really the workhorses of time and frequency, tremendous effort has been continuously done to improve performances. Quartz crystal oscillators already provide us with small, rugged, low consumption, low cost units of excellent short term stability. The main effort has to go toward decreased aging, low amplitude-frequency effect (then as a consequence better short term stability), thermal stability, low thermal transient sensitivity, and low environmental dependance (acceleration, vibration ...).

II - New concepts in resonators

For somebody who is not too much impressed by the extraordinary amount of available literature on resonators, each step of the resonator design can be reconsidered. This sometimes arises questions which, at first, look almost childish. For example :

Why do crystals have usually circular or rectangular shapes and spherical contours ?

The answer is that quartz does not care but it is usually easier to deal with those shapes or contours. If correctly handled this interrogation can lead to interesting new shapes or contours. Technical feasibility or (and) even tradition often determines the fabrication method. In addition, it is so delicate to make excellent resonators that the slightest change can be critical. This does not encourage the use of new concepts; it promotes the advantages of established procedures. But, theoretical understanding has been considerably improved

and technology has fast evolved. Also, due to an important effort in fundamental understanding of circuits, especially in N.B.S. Boulder, the concept of the oscillator itself has evolved. Improvements on the resonator alone can now clearly be seen and investigated (21), (22), (1). In other words, it is time now to make new steps in progress.

A very promising development has been the introduction of doubly rotated crystals. (3), (4) interesting for their low sensitivities to stresses in cut's plane (4) their excellent thermal behaviour including thermal transient (5) and their low amplitude frequency effect (13).

It is also possible to make drastic changes in the conventional design which obviously exhibits badly or incompletely solved boundary problems, at least to the high degre of perfection we need (a simple calculation indicates that the fact a 5 MHz crystal keeps its frequency inside the bandwidth means the thickness or equivalent thickness of resonator does not vary more than 7 Å).

Now, let us see some problems of the conventional design.

In the usual manufacturing process of a plated crystal, a thin metallic film (gold, silver, aluminium or copper) is deposited on to contoured crystal surfaces which have been previously either polished or etched. Usually the crystal is cemented to thin nickel ribbon fixations, T.C. bonded or electrobonded.

Plating should be considered first since its effects occur in the very energy trapping zone. Nevertheless the stress relaxation in the mounting structure and the contamination which is caused by the bonding process cannot be ignored.

Plating is always a very rough process for the crystal surface neighbourhood. The crystalline arrangement is partly upset, piezoelectric properties are locally modified, metallic ions penetrate inside the crystal thereby generating further ion migration. The surface of the crystal is drastically perturbed and the perturbation will not be

constant versus time so giving rise to further frequency drift. At the same time, thin film stresses cause a non negligible frequency shift, which is not stable with time. Several fundamental phenomena being involved it is difficult to predict the exact noise contribution of the plating and the exact frequency drift contribution as well. Nevertheless, it has been possible to prove that small intrinsic stresses correspond to improved aging (23) and that probably the intrinsic aging of quartz material is orders of magnitude lower than aging exhibited by plated units.

At the same time, the Q factor, mainly determined by the internal friction in quartz, is reduced by the damping due to the metal deposited on the surface. But this effect should not be exaggerated, since the Q factor obtained with plated units according to the Warner's design is close to the intrinsic Q factor of the material. Moreover excitation by reduced electrodes, annular electrodes or parallel field technique have been widely used and can be considered as a first attempt to suppress plating (24). At last, it must also be pointed out that, recently, A.G. Smagin (25) obtained very high Q factor with an experimental device using an unplated artificial crystal.

Let us consider now the frequency adjustement of plated units. It is a very important problem. Various techniques operating by additionnal deposition of metal are used, in situ environment or not.

First, the stability of the frequency adjustement, especially in situ environment, is a matter of discussion.

Second, it is generally difficult to adjust the frequency of a given unit to better than one p.p.m of the nominal frequency. (Of course, this does not include the laser machining technique which is only available for glass enclosure type units). Nevertheless a more accurate frequency adjustement is needed for some applications ; so further progress is also desirable in this domain.

New structures, all using uncoated crystals, were outlined and called B.V.A.$_n$ designs :

- if n is odd, a rather conventional bonding and a special fixation is used.

- if n is even, the design uses improved bonding and mounting.

This denomination indicates two successive steps of our attempt to reduce the crystal's noise and frequency drift contribution. Some of those new designs have already been described elsewhere (6)(7). Let us, simply review some possible solutions trying to point out some useful concepts for each given example.

1/ Usual B.V.A.$_2$ design (26) (27)

a) Evaluation -

Special emphasis is given to this design which overcomes some difficult problems caused by the conventional evaluation of piezoelectric resonators. We mainly describe here quartz material 5 MHz units but, of course, other piezoelectric materials can be used and resonators of various frequencies have been evaluated.
Our goal was to obtain an "electrodeless" resonator with a mounting exhibiting neither discontinuity nor local stress in the mounting areas. We wanted to obtain a device the frequency of which could easily be adjusted by means of a series capacitance. Then a large gap capacitance is suitable i.e. electrodes have to be located very close to the active surface of the wafer (in the micron range or even the 10 microns range). Also, and this is very important too, we planned to obtain mounting areas very accurately located and mounting means precisely known.
The B.V.A.$_2$ resonator is represented by the schemes of Fig. 1 - Fig.2.

It includes :

- a vibrating quartz crystal, ref.C, the surface of which has been very carefully prepared. The active part of the crystal is connected to the dormant part by little quartz "bridges" very precisely made and located.

- a quartz condenser made of two disks (ref. D_1 and D_2) of the same cut and orientation on which the electrodes are deposited.

- means to maintain the condenser and crystal tightly together.

- a metallic experimental enclosure which is sealed by a pinch off process.

It must be pointed out that some construction parameters, especially the support configuration parameters, can be, using this design, very precisely known. Also since the crystal is "electrodeless" and uses an all quartz structure it is very suitable for low temperature applications. Moreover, the electrodes may be deposited on insulators which have been given a curvature different from the crystal surface's. This feature gives access to additional possibilities and may be used to modify Q factor, motional parameters series resistance and frequency amplitude effect.

Such a resonator, being entirely different from a conventional resonator, needs theoretical and technical studies specially devoted to it.

The original part of the crystal evaluation will only be described here. By use of ultrasonic machining and precise lapping little bridges are left between the external dormant part of the crystal and the internal vibrating part. Those bridges have a given shape, a given thickness, a given length. The bridges can be very precisely located with respect to the thickness of the crystal (accuracy of the location : $\pm$ 10μ). Their angular position can also be very precisely

known ($\pm$ 0.04°). Of course, the technique has to be perfectly mastered (for instance, avoiding a conical ultrasonic machining is not immediate) but, with sufficient experience, the process can be considered as sure, rapid (2 or 3 minutes) and very accurate. As a consequence the middle part of the bridges can be located at the very nodes of vibration. Also, unwanted modes can be better eliminated. Since the thickness in the middle part of the bridges has ranged from 50μ to 1200μ (the usual is approximately 200μ) the bridges are not especially brittle. Any number of bridges can be left. Especially one single bridge, covering the full 360° angle, may be directly lapped so avoiding the ultrasonic machining.

It must be pointed out that the machining does not destroy the material from the cristallographic point of view.

Moreover, no additional stresses are left by the machining if the quartz wafer is subjected, prior to mounting, to annealing at about 480°C, followed by a very slight surface attack with bifluoride.

The length and thickness of the bridges have been theoretically studied. Assuming a flexure vibration of the bridge, it is found that a length of 2 mm and a thickness of 0.2 mm is a good compromise between a weak static strain and a minimum acoustical energy transmission between the vibrating and dormant part of the crystal (5 MHz fifth overtone).

The reflection of the elastic waves is not influenced by the position of the electrodes with regard to the crystal surface. It mainly depends on phenomena which occur in the boundary neighbourhood and which are due to crystalline modifications caused by machining processes and surface preparation.

The sample surface is carefully lapped and polished, so as to reduce the layer in which acoustic dissipation occurs. Defects due to machining processes are carefully investigated (X ray topography, electron microscopy and so on) so as to define the best procedures. As far as possible, we operate in a clean room atmosphere, try to process the

crystal in dry nitrogen and, of course, use the results of recent investigations for cleaning and decontamination (28)(29).
The influence of the gap has been studied. Experimentally, the Q factor is not a constant versus the total gap. The variation depends on the frequency and overtone number of the unit.
Nevertheless an investigation was started and proved that usual equivalent circuit is not sufficient.
So, starting from the exact expression of the current, we computed the Q factor versus the gap (assuming a plane infinite plate) and found a variation which gives a better account of the results (30).
Actually, a compromise must be chosen. The series resistance and the motional inductance strongly increase with the gap. Also, the frequency of the unit must be easily adjusted by a series capacitance ; so very thin gaps are suitable.
But the mechanical stability of the gap thickness is to be considered too, if ultrastable units are desired. For a 5 MHz fifth overtone we use gaps in the micron or 10 microns range. Nevertheless for resonators on the fundamental mode the gap can be larger.
Usually, the gaps are made by a special lapping process which affects the central area of D_1 and D_2 (see Fig.1). They can also be made by nickel electrodeposition. It must be pointed out that slightly different gaps can be made so giving access to very precise frequency adjustment (1 Hz for a 5 MHz fifth overtone unit).

b) Properties of B.V.A.$_2$ design -

This design has now been extensively studied for more than five years. To Summarize, it should be said that improvements of almost an order of magnitude over the conventional design seem possible for short term stability, long term drift and g sensitivity as well. At this point, the following results have already been obtained :

- stabilities $\sigma_y(\tau)$ of 5 to 6 10^{-14} over 128 s or so (2),

- drift rate of 3 to $5 \cdot 10^{-12}$ per day after one or several months,

- maximal g sensitivity of $2 \cdot 10^{-10}/g$ for an AT cut (31).

Nevertheless, the design cannot be considered as having reached its final state and some more improvements are possible. Of course, it must also be pointed out that some properties (such as thermal properties) can be totally different from the corresponding properties in a conventional unit.

2/ B.V.A.$_3$ design (27)

Since the B.V.A.$_3$ design is an improvement of the B.V.A.$_1$ design it will be very rapidly described using the scheme of Fig. 3. The vibrating quartz crystal C of a given cut, orientation, geometrical shape (in a Fig.3 a planoconvex disk) is, for instance, TC bonded (3 or even 4 bonding points ref.T) to the lower disk D_1 (which has been given a curvature identical to the curvature of the wafer's lower surface). D_1 is usually made out of quartz of the same cut and orientation. The electrodes are evaporated on the lower disk D_1 and the upper disk D_{21} or D_{22}. The upper disk is not necessarily made out of quartz and may have any radius of curvature. The intermediate ring R determines the upper gap giving access to frequency adjustment or modulation. Compared to B.V.A.$_1$ type this design has mainly the same properties but the characteristic features are greatly improved in severe environmental conditions.

3/ B.V.A. Dual Resonator (32)

The scheme of Fig.4 shows a new design comprising basically 2 B.V.A. resonators R_1 and R_2. If the axis are respectively orientated as indicated of Fig.4 (drawn for the AT cut case) very low g sensitivity

is obtained. In addition R_1 and R_2 can obviously be used in parallel or in series ; also they can have identical or different frequencies. In the parallel case the frequency difference basically does not depend on temperature if R_1 and R_2 are similar enough (32). Among other advantages this new structure gives R_1 a very precise orientation with respect to R_2 and places R_1 and R_2 in very similar thermal conditions (temperature and temperature gradients).

4/ B.V.A. design for U.H.F. range

The most important advantages of the B.V.A. design appear at high frequencies. In fact in thin or very thin coated units the "electrode effect" is relatively more important. Nevertheless, the usual B.V.A.$_2$ design is limited toward high frequencies (limitation similar to the limitation of conventional coated units). Then a new design called B.V.A. - U.H.F. has been developed for the U.H.F. range (33). Basically it comprises a very thin B.V.A. vibrating crystal in a reentrant cavity made of quartz. Resonant frequencies in the GHz range have already been obtained.

5/ Reverse structure

It is also perfectly possible to imagine quartz structures with bridge (or bridges) connecting the external vibrating to the internal dormant part (34).This structure gives excellent results with low motional resistance. The mounting in the central dormant part can be reduced to an axis and consequently can be made extremely symmetrical.

III - Superconducting cavities and crystals at low temperatures -

The analogy between the two subjects is somehow very impressive. Since stabilities down to $3 \cdot 10^{-16}$ have been demonstrated (10)for the S.C.S.O

(Superconducting Cavity Stabilized Oscillator) by Stein and Turneaure, since low temperature technology is no more a future technology this possible technique has to be considered. Both quartz crystals and superconducting cavities can be used in active oscillators or as passive reference for a free running slave oscillator (in this case all the electronic circuitry will be at room temperature). Both devices are not accurate i.e. they suffer from frequency shift and are used when the knowledge of absolute frequency is not necessary. In fact in both cases, the resonant frequency is determined by the size of the resonator. The Q currently achieved for crystal units 5 MHz are in the range of 2 to 5 10^6 at room temperature and can reach some 10^8 at helium temperature. It is to be pointed out that electrodeless resonators are the best suitable resonators for low temperatures. Superconducting cavities with Q as high as 10^{11} at 8.6 GHz have been used. Whatever are the differences similarities between superconducting cavities and crystal resonators is interesting : even similar techniques are used to evaluate the devices !(chemical polishing, oven stabilization, baking out in inert atmosphere, final preparation in glove boxes and so on). The main advantage of quartz cavities at low temperature could be their small size if it is really demonstrated that properties at helium temperature are excellent, reproducible and not affected by some hysteresis process. In this domain, the advantage of the new B.V.A. technique is obvious.

At this point and for both devices the advantages of passive and active operation can still be discussed. Basically, in active operation, the resonator (or cavity) is coupled to a negative coefficient resistor and in passive operation one can use the fact that the reflection (or transmission) coefficient of the cavity is a rapid function of frequency. A feed back type circuit is used and the high loop gain produces an oscillator whose stability is limited by the frequency fluctuations of the element (in both superconducting cavity and quartz crystal stress relaxation appear to be one important limiting

process).
Of course a passive operation will allow the reference element to be far away or at different conditions with respect to the electronics.

IV - About the use of resonant frequencies -

This section will adress some basic problems related to the use of resonators giving some references for specific problems of circuits. Actually, the traditional concept of the best quartz oscillators has recently been reconsidered especially in N.B.S. Boulder (Time and frequency standards section). This effort, mainly due to H. Hellwig, S.R. Stein and F.L. Walls, was initiated with passive measurements on crystals in 1974 (21). Actually, passive measurements by Walls and al (on an initial suggestion of D. Halford) demonstrated that the intrinsic noise of the best quartz crystal resonators is significantly less than the noise observed in oscillators employing those resonators. It has been possible to predict the performances of the composite system based on the measured performances of its components (1). Also it has been possible to compare various types of resonators (22), in various conditions, from the noise point of view. In fact one has come to the point where circuit stability can be distinguisted from crystal stability. Since superior performances have been demonstrated with a passive reference crystal oscillator (2) we can now determine which problems come from the crystal and which problems are due to the circuit. At this point, we can say that at least some crystals are more stable than others and that it is not obvious to take advantage of superior capabilities of given resonators. It is certainly important to make more investigation on passive and active circuits, assuming the main features and capabilities of the crystal are known by passive methods. Nevertheless, it is still too soon to recommend either passive methods or active ones, at least for any integration time. Anyway, the conclusion is that important work should be performed on cir-

cuits. Also, crystal resonators or cavities should be designed taking into account the knowledge of circuits (the inverse is, of course, also true). It is interesting to point out that apparent asymptotic limits of performances have been observed for one or two decades (some Xtal oscillators, built many years ago, are still among the best ones). Is it completely ridiculous to imagine that the limitation due to Xtal resonators is somehow of the same order of magnitude as the limitation due to electronics ? If it was true, the proper answer should be to use at the same time better electronics and better crystal units.

V - Discussion of recent advances, perspective for near future and conclusions -

The recent results obtained by the S.C.S.O. (Superconducting Cavity Stabilized Oscillator) are the most inpressive ($3 \cdot 10^{-16}$ noise floor achievement) and provide us with a strong stimulation for superior performances.

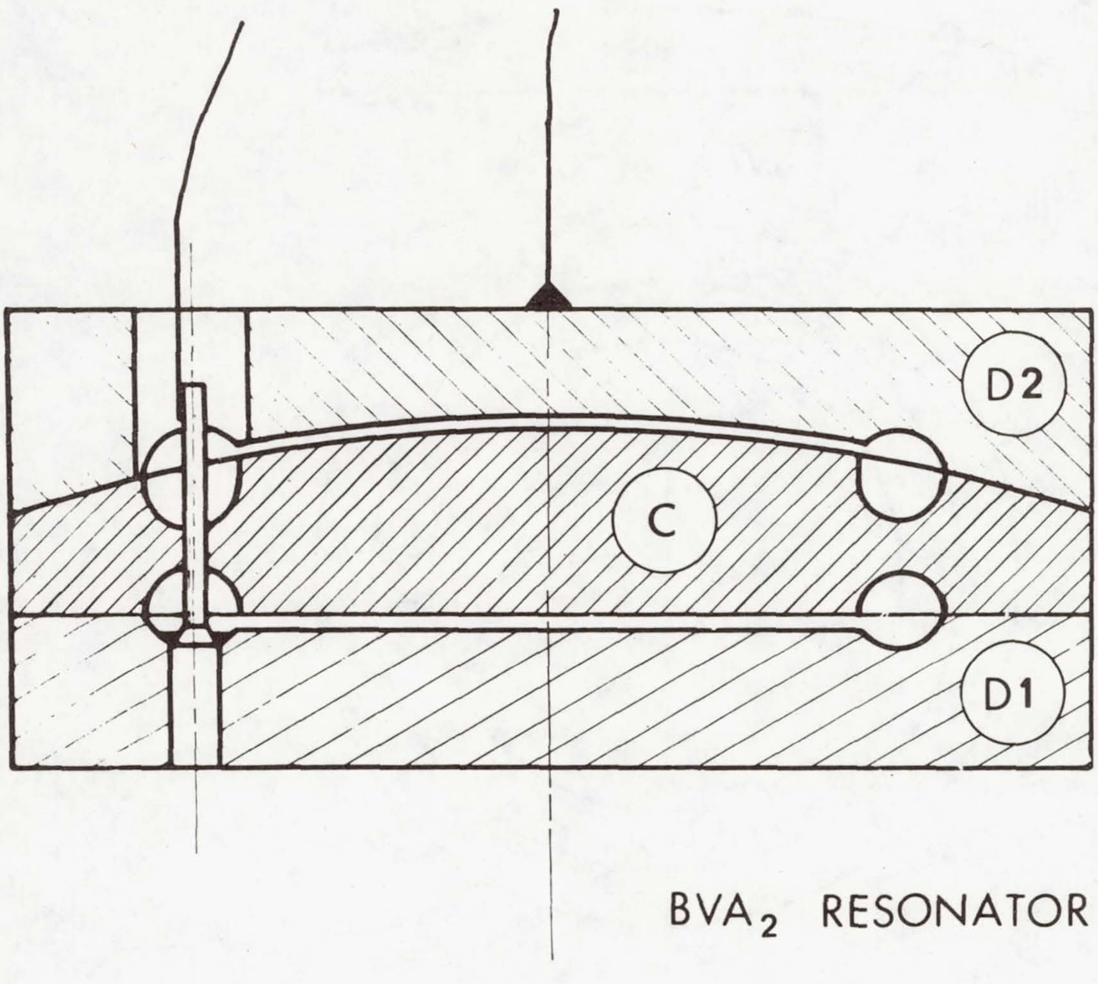
From what has been seen, realistic performances for quartz oscillators (at room temperature) can be expected to be available in the near future. Drifts on the order of 10^{-12}/day, short term stabilities in the 10^{-14} range, acceleration sensitivities of $2 \cdot 10^{-10}$/g (AT cut) and $2 \cdot 10^{-11}$/g (doubly rotated cut), acceleration sensitivity compensation, very low noise capabilities and excellent spectral purity already seem possible.

Nevertheless various investigations in domains ranging from environmental influence (35)(36)(37)(38)(39)(40)(41)... quartz resonator design,measurement and application (30)(42)(17)(43)(44) ... resonator processing (45)(46)(47)(48)(49)(50) ... fundamental properties of materials (51)(52) ... oscillator design specification and measurement (2)(53)(54)(55) ... are still currently being done and applied. Further advances are still possible on each individual feature ; also it can be useful (and not obvious) to obtain excellent performances at

the same time for various features (including environmental sensitivity).
In addition, quartz crystals seem very promising at low temperature for oscillator applications, since very high Q factors are possible especially with the "electrodeless technique". Of course, quartz resonators can also be used (especially doubly rotated cuts) to make very convenient temperature sensors. In this domain the main effort has to go toward very low hysteresis effects and high reproducibility.
Though it has not really been demonstrated yet, the B.V.A. technique should be of interest for this purpose.
Furthermore, it is to be pointed out that the concept of composite oscillator systems (56) or composite resonator systems (32) is a very powerful tool to meet the severe requirements which will be needed in near future.

Acknowledgements -

The author, currently guest worker in "Time and Frequency Standards Section" N.B.S. Boulder, would like to thank especially Drs H. Hellwig, S.R. Stein, F.L. Walls and C.M. Manney for many helpfull discussions. He highly appreciates the opportunity that was given to him to work in this laboratory environment and would like to thank the organizations that made it possible (in France D.R.E.T. and Ministry of Universities, in the U.S. National Bureau of Standards, U.S. Department of Commerce).



BVA_2 RESONATOR

FIG - 1

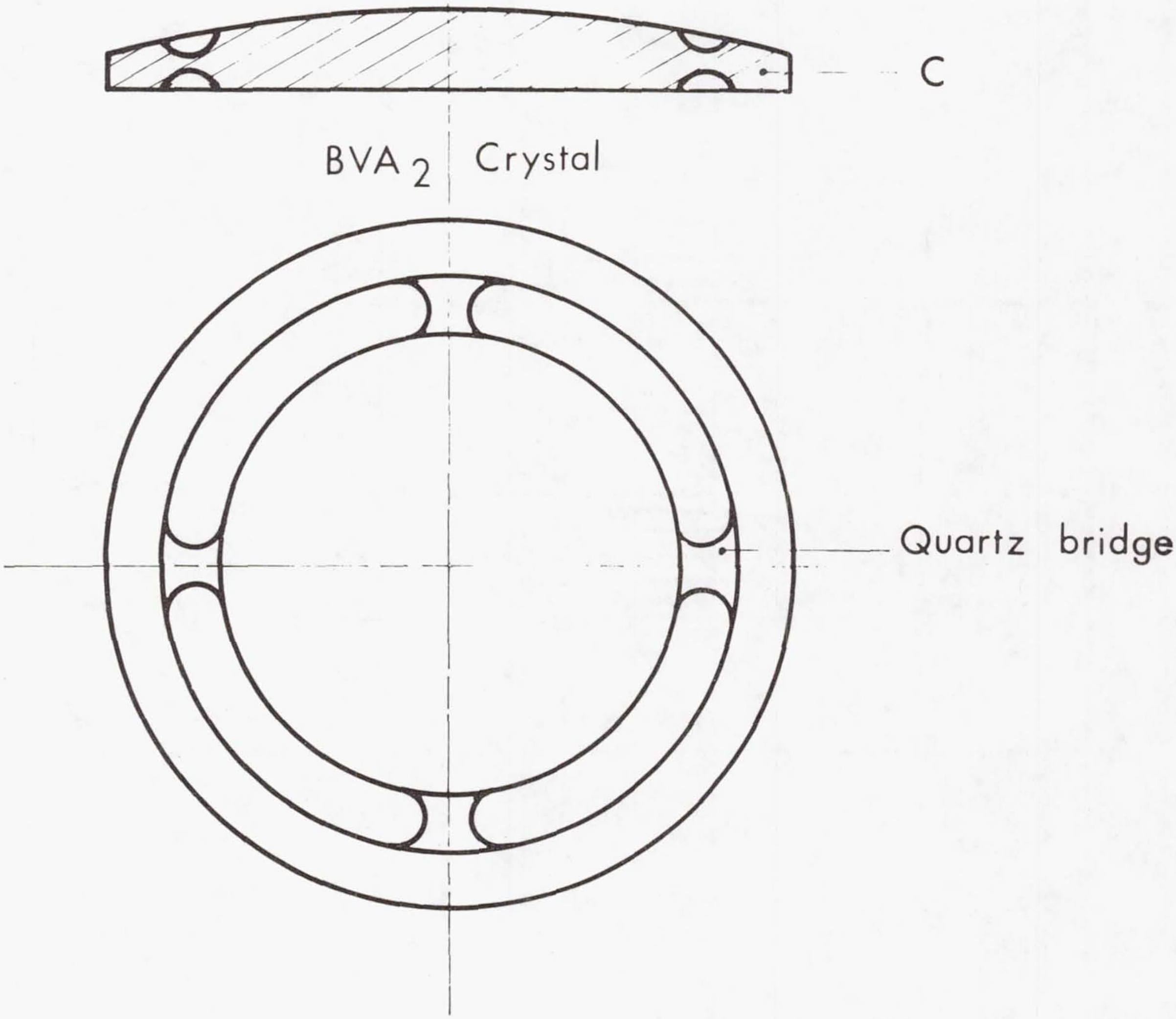
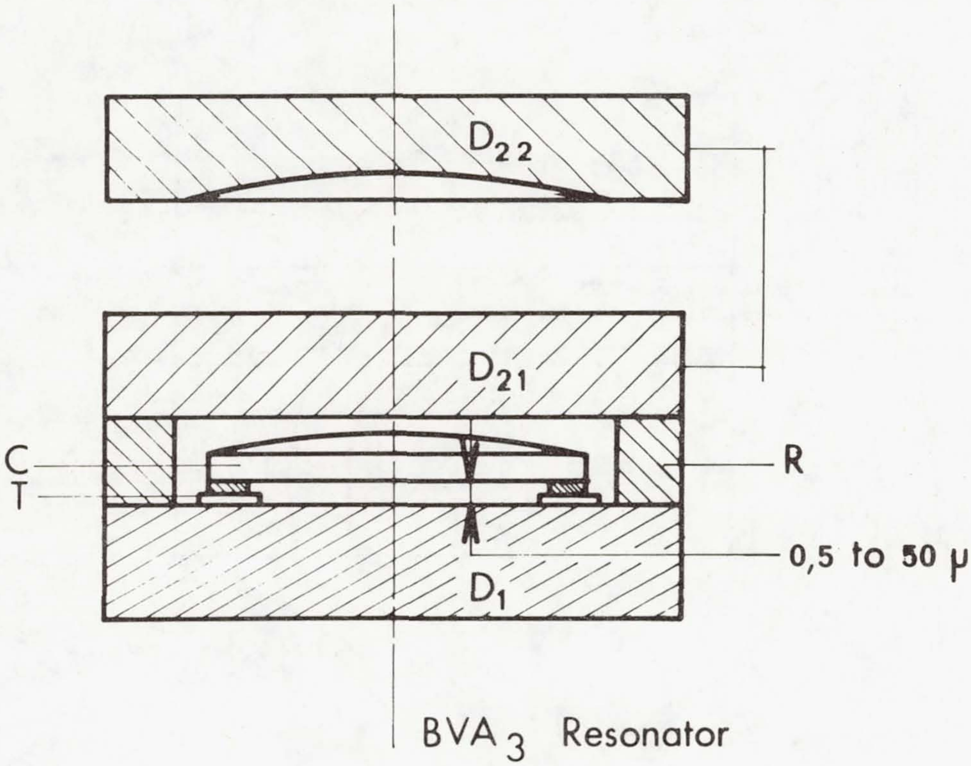


FIG-2



BVA 3 Resonator

FIG 3

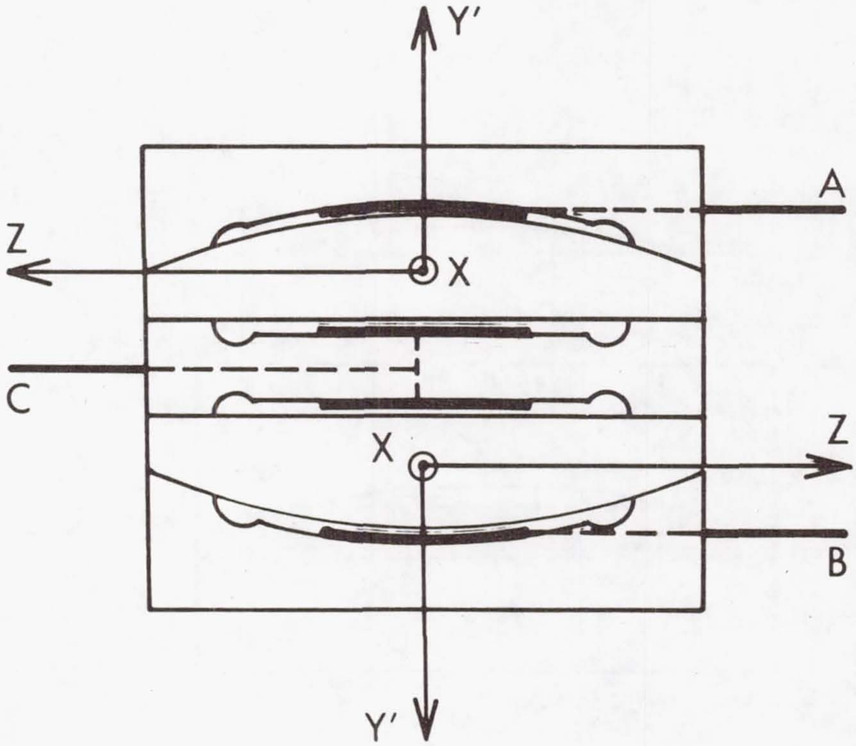


FIG 4

References -

(1) - F.L. Walls - S.R. Stein -
"A frequency lock system of improved quartz crystal oscillator performance" -
I.E.E.E. Trans. Instrum. Meas. - IM - 27, 249 - 252 (1978).

(2) - S.R. Stein - C.M. Manney,Jr, F.L. Walls, J.E. Gray - N.B.S. and R. J. Besson E.N.S.C.M.B. -
"A systems approach to high performance oscillators"
Proc. 32nd A.F.C.S. - pp. 527 - 530 (1978)

(3) - A. Ballato -
"Doubly rotated thickness mode plate vibrators"
Physical Acoustics - Vol. XIII - pp. 115-181 - Ac. Press NY. (1978)

(4) - E.P. Eer Nisse -
"Quartz Resonator Frequency shifts arising from electrode stress"
Proc. 29th A.F.C.S. - pp. 1 - 4 (1975) and "Calculations on the stress compensated (S.C.-cut) quartz resonator" - Proc. 30th A.F.C.S. pp 8 - 11 (1976)

(5) - J. Kusters -
"Transient thermal compensation for quartz resonators"
I.E.E.E. Trans. Sonics and Ultrasonics SU-23 pp-273 - 276 (1976)

(6) - R.J. Besson -
"A new piezoelectric resonator design"
Proc. 30th - A.F.C.S. - pp. 78 - 83 (1976)

(7) - R.J. Besson -
"A new electrodeless resonator design"
Proc. 31rst - A.F.C.S. pp - 147 - 152 (1977)

(8) - S.R. Stein -
"Application of superconductivity to precision oscillators"
Proc. 29th - A.F.C.S. - pp. 321 - 327.

(9) - J.P. Turneaure in Proc. 1972 - Applied Superconductivity Conference (I.E.E.E. - New York 1972) p. 621.

(10) - S.R. Stein and J.P. Turneaure in Future Trends in Superconductive Electronics (A.I.P. New York - 1978) pp. 192 - 213.

(11) - E. Hafner -
"Crystal resonators"
I.E.E.E. Trans. Sonics and Ultrasonics - Vol. SU 21 N°4
pp. 220 - 237 (October 1974)

(12) - A. Ballato and G.J. Iafrate -
"The angular dependence of piezoelectric plate frequencies and their temperature coefficients"
Proc. 30th - A.F.C.S. - pp. 141 - 156 (1976)

(13) - J.J. Gagnepain - J.C. Ponçot and C. Pegeot -
"Amplitude - Frequency behaviour of doubly rotated quartz resonators"
Proc. 31rst A.F.C.S. - pp. 17 - 22 - (1977)

(14) - J.J. Gagnepain - R.J. Besson -
"Non linear effects in piezoelectric quartz crystals"
Physical Acoustics - Vol. XI - Ac. Press - N.Y. (1975)

(15) - D. Hauden - M. Michel - J.J. Gagnepain -
"Higher order temperature coefficients of quartz S.A.W. Oscillators"
Proc. 32nd A.F.C.S. - pp. 77 - 78 - (1978)

(16) - J.H. Balbi - J.A. Duffaud and R.J. Besson -
"A new non linear analysis method and its application to quartz crystal resonator problems"
Proc. 32nd - A.F.C.S. - pp. 162 - 168 - (1978)

(17) - H.F. Tiersten - R.C. Smythe -
"An analysis of overtone modes in contoured crystal resonators"
Proc. 31st A.F.C.S. - pp. 44 - 47 - (1977)

(18) - H.F. Tiersten -
"Linear piezoelectric plate vibrations"
Plenum Press - (1969)

(19) - R.D. Mindlin -
"An introduction to the mathematical theory of vibrations of elastic plates"
US Army Signal Corps - Fort Monmouth - N.J. 1955

R.D. Mindlin - W.J. Spencer -
"Anharmonic thickness twist overtones of thickness shear and flexural vibrations of rectangular AT cut plates"
J. Acoust. Soc. Am 42 p. 1268 (1977)

(20) - H.F. Tiersten -
"Analysis of non linear resonance in rotated Y cut quartz thickness shear resonators"
Proc. 29th A.F.C.S. - pp. 49 - 53 (1975)

(21) - A.E. Wainwright - F.L. Walls - W.D. Mc Caa -
"Direct measurement of the inherent Frequency stability of quartz crystal resonators"
Proc. 28th - A.F.C.S. - pp. 117 - 180 (1974)

(22) - J.J. Gagnepain -
"Fundamental noise studies of quartz crystal resonators"
Proc. 30th A.F.C.S. - pp. 84 - 91 (1976)

(23) - E. Hafner - R.S. Blewer -
Proc. I.E.E.E. 56 336 (1968)

(24) - R. Bechmann - Proc. A.F.C.S. (1953)
and W. Ianouchevsky - Proc. 17th - A.F.C.S. pp. 233-247 (1963)

(25) - A.G. Smagin -
Pribory i Tecknika Ekspeimenta N°6 - pp. 143 - 145
(Nov. Dec. 1974)

(26) - French Patents n° 76 01 035 (2.338.607) n° 76 16 289-
"Résonateur à quartz à électrodes non adhérentes au cristal"

(27) - French Patent n° 78 02 261 -
"Résonateur piézoélectrique à cristal auto-suspendu"

(28) - J.R. Vig - C.F. Cook - K. Schwidtal - J.W. Lebus - E. Hafner -
"Surface studies for quartz resonators"
Proc. A.F.C.S. - pp. 96 - 108 (1974)

(29) - J.R. Vig - J.W. Lebus -
"Further results on UV cleaning and Ni electrobonding"
Proc. A.F.C.S. - pp. 220 - 229 (1975)

(30) - B. Dulmet -
"Etude théorique et expérimentale de quelques paramètres régissant le fonctionnement des résonateurs à quartz"
Thesis n°87 - Besançon 1978

(31) - Measurements performed in O.N.E.R.A. - 92230 Chatillon - France

(32) - French Patent n° 78 28 728 -
"Bi-résonateur piézoélectrique"

(33) - French Patent n°78 02 261 -
"Résonateur piézoélectrique à cristal autosuspendu"

(34) - French Patent n°77 17 309 -
"Résonateur piézoélectrique à suspension amélioré"

(35) - M. Valdois - J. Besson and JJ Gagnepain -
"Influence of environment conditions on a quartz resonator"
Proc. 28th A.F.C.S. - pp. 19 - 32 (1974)

(36) - D. Janiaud -
"Modélisation de l'influence d'une accélération sur la fréquence des résonateurs à quartz" - Besançon 1978.

(37) - P C Y Lee and Kuang-Ming Wu -
"The influence of support configuration on the acceleration sensitivity of quartz resonator plates"
Proc. 31st - A.F.C.S. - pp. 29 - 34 (1977)

and "Effects of acceleration on the resonance frequencies of crystal plates"
Proc. 30th - A.F.C.S. - pp. 1 - 7 (1976)

(38) - A. Ballato - T. Lukaszek - E.P. Eer Nisse -
"The force frequency effect in Doubly Rotated quartz Resonators"
Proc. 31st - A.F.C.S. - pp. 8 - 16.

(39) - H.F. Tiersten and B.K. Sinha -
"Temperature induced Frequency changes in Electroded AT-cut Quartz Thickness shear Resonators"
Proc. 31st - A.F.C.S. - pp. 23 - 28.

(40) - A. Ballato - J.R. Vig -
"Static and dynamic Frequency-Temperature behavior of singly and doubly rotated, oven controlled quartz resonators"
Proc. 32nd - A.F.C.S. - pp.180 - 188.

(41) - Measurements of environmental influence on B.V.A. crystals - Work currently performed on a cooperation basis by O.N.E.R.A. and E.N.S.C.M.B. France.

(42) - R.J. Besson - B.M. Dulmet and D.P. Gillet -
"Analysis of the influence of some parameters in quartz bulk resonators"
Proc. Ultrasonics Symposium (1978).

(43) - E. Hafner - W.J. Riley -
"Implementation of bridge measurement techniques for quartz crystal parameters"
Proc. 30th - A.F.C.S. - pp. 92 - 102 (1976)

(44) - A. Ballato - R. Tilton -
"Ovenless Activity Dip Tester"
Proc. 31st - A.F.C.S. - pp. 102-107 (1977)

(45) - J.R. Vig - J.W. Lebus - R.L. Filler -
"Chemically polished quartz"
Proc. 31st - A.F.C.S. - pp. 131 - 143 (1977)

(46) - M. Berté -
"Acoustic bulk wave resonators and filters operating in the fundamental mode at frequencies greater than 100 MHz"
Proc. 31st - A.F.C.S. - pp. 122 - 125 (1977)

(47) - P.D. Wilcox - G.S. Snow - E. Hafner - J.R. Vig -
"A new ceramic flat pack for quartz resonators"
Proc. 29th - A.F.C.S. - pp. 202 - 210 (1975)

(48) - R.N. Castellano - J.L. Hokanson -
"A survey of Ion beam milling techniques for piezoelectric device fabrication"
Proc. 29th - A.F.C.S. - pp. 128 - 134 (1975)

(49) - M.L. White -
"Clean surface technology"
Proc. 27th - A.F.C.S. - pp. 89 - 94 (1973)

(50) - R.L. Filler - J.M. Frank - R.D. Peters - J.R. Vig -
"Polyimide bonded resonators"
Proc. 32nd - A.F.C.S. - pp. 290 - 298 (1978)

(51) - C.K. Hruska -
"The electroelastic tensor and other second order phenomena in quasi linear interpretation of the polarizing effect with thickness vibration of α quartz plates"
Proc. 31st - A.F.C.S. - pp. 159 - 170 (1977)

(52) - F. Euler - P. Ligor - A. Kahan - P. Pellegrini - T. Flanagan - T. Wrobel -
"Steady state radiation effects in precision quartz resonators"
Proc. 32nd - A.F.C.S. (1978)

(53) - F.L. Walls - S.R. Stein -
"Accurate measurements of spectral density of phase noise in

devices"
Proc. 31st - A.F.C.S. - pp. 335 - 343 - (1977)

(54) - D.W. Allan -
"Measurement of frequency and frequency stability of precision oscillators "
N.B.S Technical Note 669 (1975)

(55) - J. Rutman -
"Characterization of phase and frequency instabilities in precision frequency sources"
Proc. I.E.E.E. - Vol. 66 N°9 - pp. 1048 - 1075 (Sept. 1978)

(56) - R.S. Stein - F.L. Walls -
"Composite oscillator systems for meeting user needs for time and frequency" somewhere in these proceedings.

QUESTIONS AND ANSWERS

DR. HELMUT HELLWIG, National Bureau of Standards

I would like to make a comment before I ask for questions. As far as I can remember, the original suggestion at NBS to build a measurement system to measure passive crystal resonators is due to Donald Halford. Are there any questions?

MR. JOHN R. VIG, Army Electronics R&D

You mentioned that for your dual BVA design, the acceleration sensitivity was a factor of 20 times better. Better than what? You mentioned several sensitivities before that, and I wasn't sure which.

DR. BESSON:

Okay. Better than the usual for an AT cut, which is 8×10^{-11}.

MR. VIG:

Have you tried the dual for the SC?

DR. BESSON:

Not yet. It is done but not tested.

DR. CARL HOSHKA, York University

How long does it take to produce one of these crystals from a blank?

DR. BESSON:

Well, some steps are faster; some are longer. The step which is critical, and that I did not mention, is a step of addressing the frequency. And with those crystals, you do have a great facility to make the frequency adjustment because you have only to change the gap.

And then you have only to evaluate, let's say, condensers with different gaps, and it comes out that you just have to change gaps, and that is it. So I would talk about prices, okay? I think prices so far as we see in the investigation lab, which is not industry, are 1.2 to 1.7 for a usual crystal.

DR. HOSHKA:

And time-wise? If we were to concentrate just on the mechanical part, the crude part, cutting?

DR. BESSON:

It is not that different, finally.

DR. MICHEL TETU, Laval University:

In the dual crystal resonator that you have, you showed that the central electrode was coated on the center quartz. Does this mean that it will affect the long-term stability of the oscillator--drift and things like this?

DR. BESSON:

I don't understand. The central piece of quartz is not vibrating.

DR. TETU:

No, but you have the electrode on it, right?

DR. BESSON:

Yes, and it doesn't make any difference. The two BVA designs, basically, have no electrode on the vibrating crystal. That (the coated central piece of quartz) is simply a common electrode, so--No, basically you have the same design.

MR. VIG:

Have any manufacturers expressed an interest in making BVA crystals? And if so, do you expect them to become commercially available in the future?

DR. BESSON:

Yes, I think the crystals will be made (manufactured) but I can't say when.

DR. HELLWIG:

I think Dr. Besson has a whole series of patents, and I think it is reasonably public knowledge that some companies are interested in rights, and have secured rights to those patents.

DR. ALFRED KAHAN, Rome Air Development Center

It seems that the ideal resonator would be a double-rotated SC cut, or some type cut, electrodeless, operating below liquid helium at Lambda temperatures. Have you put those three items together?

DR. BESSON:

We are currently doing it at NBS, but we don't have any results yet.

SHIELDING OF LONGITUDINAL MAGNETIC FIELDS WITH THIN, CLOSELY, SPACED, CONCENTRIC CYLINDRICAL SHELLS WITH APPLICATIONS TO ATOMIC CLOCKS

S.A. Wolf, D.U. Gubser, and J.E. Cox
Naval Research Laboratory, Washington, DC 20375

ABSTRACT

Formulae for the longitudinal shielding effectiveness of thin, closely, spaced concentric cylindrical shells have been developed and experimentally tested. For shields which cannot be oriented, or which change their orientation in the ambient field, the shielding effectiveness for longitudinal fields is generally the limiting criteria and no design formulae have been presented for more than two shields. In this paper a general formula is given for the longitudinal shielding effectiveness of N closed concentric cylinders. The use of these equations is demonstrated by application to the design of magnetic shields for hydrogen maser atomic clocks. Examples of design tradeoffs such as size, weight, and material thickness will also be discussed.

Experimental results on three sets of shields fabricated by three manufacturers have been obtained. Two of the sets were designed employing the techniques described above. Agreement between the experimental results and the design calculations is then demonstrated.

INTRODUCTION

Shielding of magnetic fields is very important for the stable operation of atomic clocks. In the case of a hydrogen maser the requirements for shielding the cavity are quite stringent ($\Delta H \lesssim 10^{-10}$ tesla). Furthermore, for possible spaceborne applications, size and weight become added constraints. For these spaceborne applications a reliable method is required to accurately estimate the shielding effectiveness (ratio of internal to applied magnetic field) of concentric shields so that a design minimizing the size and weight of a shield set can be specified.

Formulae for shielding effectiveness, of open ended concentric cylindrical shells of high permeability material

in a transverse magnetic field are readily available.[1-4] However, no general formulae exists for shielding longitudinal fields although Mager[4-6] has given a relationship between transverse and longitudinal shielding effectiveness for 1 cylinder (with and without end caps) and estimated a relationship for two open concentric cylinders. For shields which cannot be oriented, or which change their orientation in the ambient field, shielding effectiveness for longitudinal fields is generally the limiting criteria.

The general equation for the longitudinal shielding effectiveness G_N of N thin, closely spaced, high permeability cylindrical shields is given by[7]

$$G_N^{\ell} = \frac{1}{2}\,(u_{i+1} + v_{i+1}) \tag{1}$$

where

$$u_{i+1} = (1 + g_i^{\ell} s_{i,i+1}^{\ell})u_i + s_{i,i+1}^{\ell} v_i$$

$$v_{i+1} = g_i^{\ell} u_i + v_i$$

with

$$u_1 = v_1 = 1$$

$$s_{N,N+1}^{\ell} = 0$$

and

$$g_i^{\ell} = \frac{\mu_i^{\ell} t}{b_i} = \frac{4D\mu_i t}{B_i} \times \begin{cases} 1 & \text{(open shields)} \\ [1+b/L]^{-1} & \text{(closed shields)} \end{cases}$$

$$s_{i,i+1}^{\ell} \approx \frac{3(b_{i+1}-b_i)}{4b_{i+1}} \times \begin{cases} [1+b/L]^{-1} & \text{(open shields)} \\ 1 & \text{(closed shields)} \end{cases}$$

Here, μ_i is the permeability of the i^{th} shield, t is the thickness of the individual shields, b_i is the radius of the i^{th} shield, D is the demagnetization factor of the cylinder, L and b are the average length and radii of the shield set, and u_i and v_i are symbols used to generate a recursive relationship. This formula is valid when

$t/b_i \ll 1$, $s_{i,i+1} \ll 1$, $g_i^{\ell} \gg 1$, and $1 < L/b < 8$. The development of this equation, an easy method of pictorially representing the many terms in the recursive equation 1, and generalization to arbitrarily large L/b ratios are found in reference 7.

DESIGN CONSIDERATIONS

A shield set designed for a spaceborne hydrogen maser must provide shielding over the entire cavity such that the changing external field will not perturb the internal field to the extent that the consequent frequency shift will be outside the specifications for the maser. In this section the various parameters that determine the overall shielding factor will be discussed in the framework of design for an optimized shield set.

Design parameters in Eqs. 1 are the shield thickness t, shield spacing $s_{i,i+1}$ and shield number N. The inner radius b_1 is usually set by the particular application and the permeability is a property of the material. Optimum shield spacing is set by the condition

$$\mu_i^{\ell} \frac{s_{i,i+1}}{b_i} = \mu_j^{\ell} \frac{s_{j,j+1}}{b_j}$$

however, for closely spaced shields, such optimization gives slight improvement over equally spaced shields, and is not a significant design consideration in most instances.

More important options are the choices between N, t, and b_N consistent with a required shielding effectiveness. Equation 1 has been used to generate a set of t and b_N values for N equally spaced closed shields with length $L \approx 3b_N$. The results are shown in fig. 1. Choices between b_N and t will depend on the physical limits imposed by a particular shielding application.

Often it is important to minimize the weight (or equivalently the amount of material) necessary for a given shielding requirement i.e. a shielding effectiveness criteria. The weight of a set of closed cylindrical shields is

$$W_N \approx 2\Pi t \rho \sum_{i=1}^{N} (b_i L_i + b_i^2)$$

where ρ is the density of the shielding material. For $L_i \approx 3b_i$ this equation reduces to

$$W_N \approx 8\Pi t \rho \sum_{i=1}^{N} b_i^2$$

Using this equation, the weight was calculated for N equally spaced shields as a function of t and b_N for a given shielding effectiveness. Figure 2 shows such a calculaton for 4 and 5 equally spaced shields where the shielding effectiveness criteria are indicated. It is noted that an optimum b_N exists for minimizing the weight. One can make shields physically smaller by reducing b_N but the thickness of the shields increases rapidly and the weight goes up. Likewise, the weight increases if b_N increases beyond the optimum value since the large b_N values more than offset the reduced t values.

For shielding to very low magnetic fields (10^{-10} tesla) the initial permeability μ_o of the material is an important material parameter, especially for the innermost shields. Since the permeability is a function of the internal induction field B inside the material, the value of μ will increase in the outer shields. For high permeability alloys, μ typically has a maximum μ_m near $B \approx 2000$ gauss which is more than 10 times μ_o. Induction B inside a cylindrical shell is approximately given by

$$B \text{ (gauss)} \approx (5/2)\left(\frac{bH_o}{t}\right)$$

where H_o is the field outside the shield. This relation along with the manufacturers published μ(B) curve, should be used to estimate μ of the outermost shield. Additional optimization can be obtained if the thickness t is selected to achieve μ_m in the outer shield.

The axial magnetic field profile within a partially closed cylindrical set of shields is usually dominated by the exponential decay of the external field as it enters the shielded region through the holes.[7]

$$H(x) \approx 1.1\left(\frac{L}{b_1}\right)^{1/2} \exp\{-k(0.75L-x)/2r\}+H_i \qquad (2)$$

where x is the axial distance from the center, H_o is the external field, L is the average length of the shields, 2r is the

hole diameter in the end caps, and H_i is the internal field far from the entrance holes as determined by the shielding effectiveness, G_N. The profiles measured on two differently dimensioned magnetic shields and 3 different hole diameters were consistently fit to this equation using the theoretical value for k of 2.26. This equation can therefore be used to establish the minimum length of closed magnetic shields with access holes in order to maintain a specified shielding effectiveness for a given axial distance near the center of the shields.

SHIELD ACQUISITION AND EXPERIMENT TESTS

Using the above considerations, a set of shields was designed for the NRL passive hydrogen maser. The design shielding factor using the manufacturer's value of permeability was 6×10^5 over a centrally located 5" long region in the shield. This shielding would provide a more than adequate safety margin to insure that the maser's frequency stability specification of 1 part in 10^{14} would not be compromised by an external field change of $\pm 10^{-4}$ T ($\pm$ 1 gauss). Several shield sets conforming to the final design were purchased from two different manufacturers. In addition, a larger shield for an SAO VLG-11 ground based maser was purchased from a third manufacturer. Figure 3 shows the schematic design of the NRL designed shields along with an actual photograph of one set. The manufacturers' quoted values were approximately identical. Specifications of the shields including dimensions and manufacturer are shown in Table I. Shield set 3 was significantly larger than sets 1 and 2, the end caps were hemispherical instead of conical, and the entrance ports in the ends of the shields were different in size with no flared extensions.

Magnetic measurements of the shields were made in an 11.3 meter diameter Braunbek coil system at the Spacecraft Magnetic Field Site, Goddard Space Flight Center, Greenbelt, Maryland. (see figure 4). This coil system actively compensates for changes in the earth's magnetic field and is capable of nulling the earth's field to better than 1 nT over a 1.3 meter sphere. In addition, the system can apply a field, known to an accuracy of 1 nT, over this volume with a magnitude as large as 60 μT. Shielding effectiveness was determined by incrementally increasing the field in a specified direction while monitoring the internal field. Both longitudinal and transverse axial shielding effectiveness were determined using a fluxgate magnetometer with a 0.1 nT resolution. Prior to each measurement, the shields were depermed to a remanant internal field less than 1 nT.[8]

In set 1, each individual shield was measured in order to arrive at the experimental μ_i^ℓ value. Next the shields were sequentially assembled and measured - providing shielding effectiveness's for 1,2,3, and 4 nested sets. These measurements, shown in Table I, were used to verify the equation for G_N^ℓ. Also, shown in parenthesis for set 1 are the μ_i^ℓ values estimated using manufacturers published specifications. The G_4^ℓ calculations using these estimated μ_i^ℓ values are too high by a factor of 2.

For the second set of shields, μ was experimentally determined at both high and low induction values for only one shield, and the results were assumed to apply for the remainder of the shields (i.e. the shields were not checked for material variability). Equation 1 was then used to calculate the measured G_4^ℓ and the result was experimentally verified (Table I). Again estimated μ_i^ℓ values for these shields using manufacturers specifications are shown in parenthesis. For this set, the calculated G_4^ℓ using estimated μ_i^ℓ values are almost an order of magnitude too high.

For the 3rd set of shields only the manufacturers values were used. The calculated G_4^ℓ value is almost double the experimental value.

It has been our experience that the $\mu(B)$ value supplied by the manufacturer is an upper limit that is not practically obtained in fabricated shields. The $\mu(B)$ plots are generated by measurements with a permeammeter on a small test piece, rather than on a fabricated cylinder in a uniform magnetic field. It is not surprising that the permeabilities determined by measuring the shielding of cylinders in uniform fields are lower than predicted on the basis of the manufacturers graphs. This must be taken into account when designing a shield set either by measuring the μ of a cylinder or by adding an adequate safety margin to the design calculations.

Figure 5 shows the measured axial field profiles for shields 1 and 3. The solid lines represent Eq. 2 for the appropriate shield length and hole diameter with k = 2.26.

CONCLUSION

Formulae presented here for the shielding effectiveness and the field profile of a closed set of N concentric cylindrical shields with access holes can be readily used to design and optimize magnetic shields for specific applications. Such design considerations have been successfully employed in the development of magnetic shields for hydrogen masers. The largest uncertainty in designing magnetic shields relates to the variability of quoted μ values. The only certain way to obtain reliable values for precise shielding calculations is to actually measure μ for at least one shield.

ACKNOWLEDGEMENTS

We thank the staff of Goddard's Spacecraft Magnetic Test Site for their generous assistence in performing the experimental tests. We also appreciate the encouragement and support given to us by Dr. V.J. Folen all through the course of this work. We also thank Roger Easton and Al Bartholomew for their generous support of this work. The Naval Electronics Systems Command is also acknowledged for their support of this work.

REFERENCES

1. T.E. Sterne, Rev. Sci. Instr. 6, 324 (1935).

2. W.G. Wadey, Rev. Sci. Instr. 27, 910 (1956).

3. S.M. Freake and T.L. Thorp, Rev. Sci. Instr. 42, 1411 (1971).

4. A.J. Mager, IEEE Trans. Mag. MAG-6, 67 (1970).

5. A. Mager, J. Appl. Phys. 39, 1914 (1968).

6. A. Mager, Z. Angew. Phys. 23, 381 (1967).

7. D.U. Gubser, S.A. Wolf, and J.E. Cox, Rev. Sci. Instr. to be pubished.

8. S.A. Wolf and J.E. Cox, NASA Tech. Memo. 78104, p 371 March, 1978.

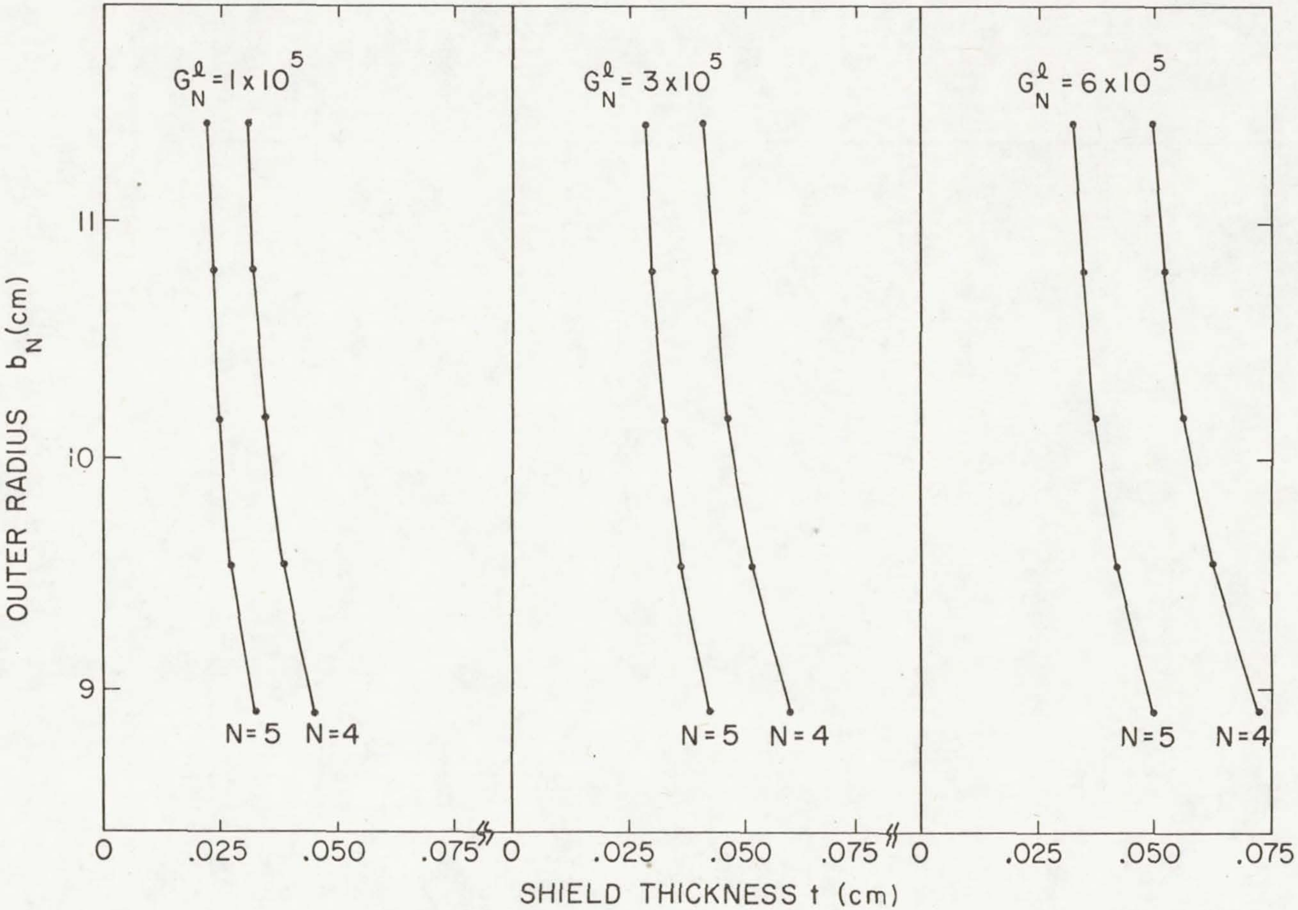


Fig. 1. Variation of outer shield radius and shield thickness for various G_N values.

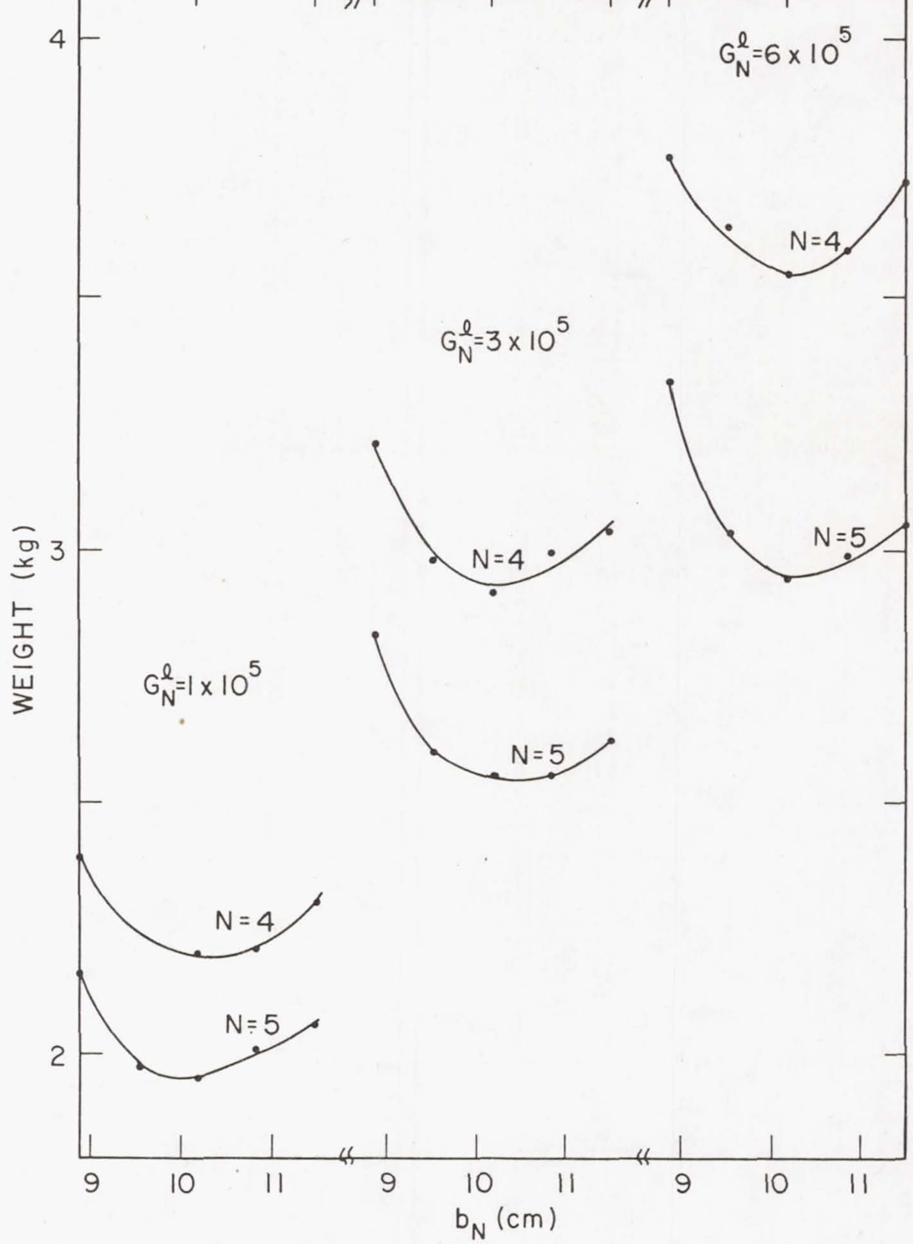


Fig. 2. Variation of total shield weight as a function of outer shield diameter for various G_N values.

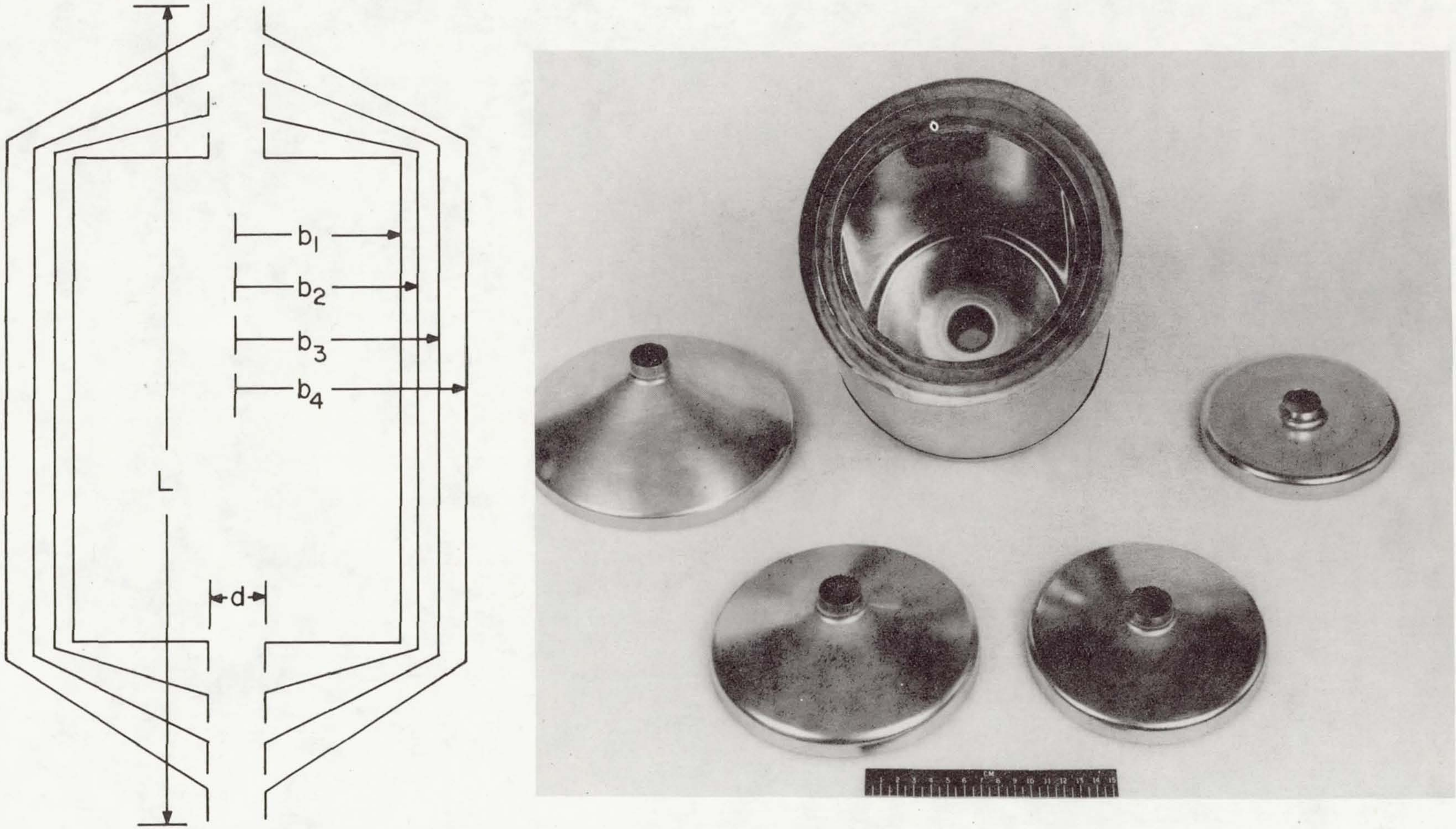


Fig. 3a. A schematic design of the two smaller sets of shields.

b. Photograph of a nested set of shields and their end caps.

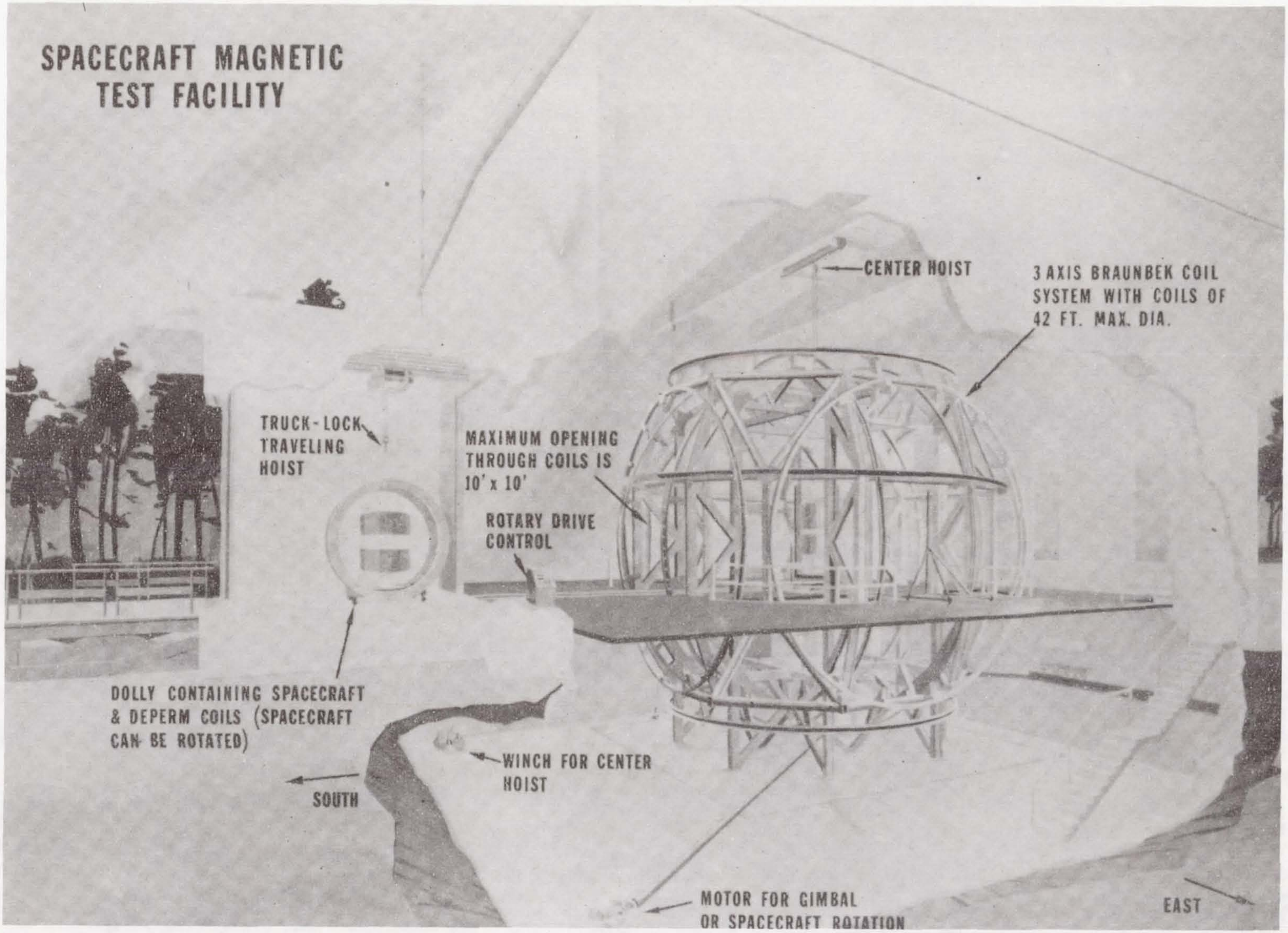


Fig. 4. Braunbek coil system at the Spacecraft Magnetic Field Site, Goddard Space Flight Center, Greenbelt, MD.

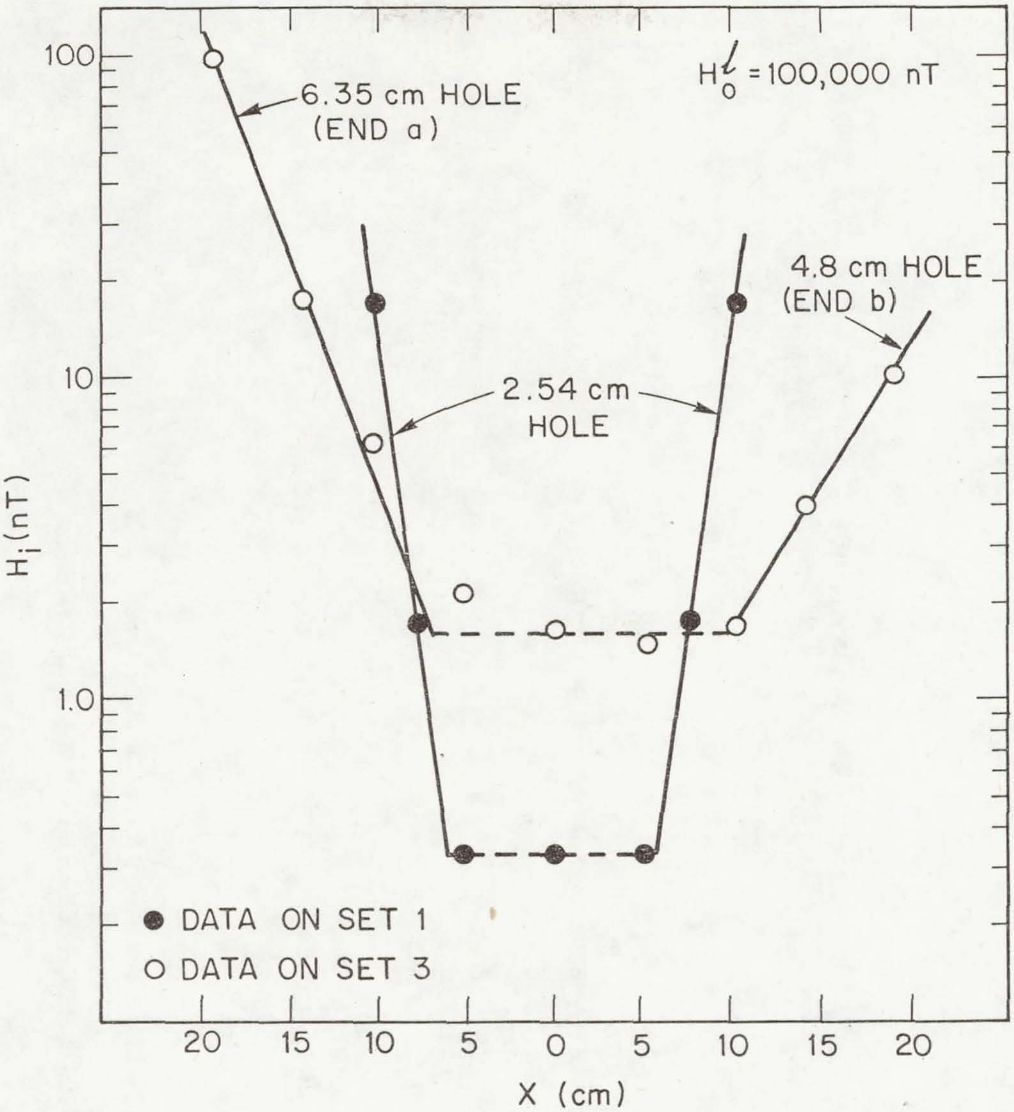


Fig. 5. Axial field profiles for shield sets 1 and 3 (Table I).

Table I Shield Parameters

	RADIUS (cm)				PERMEABILITY ($\times 10^{-4}$)				G_n^ℓ ($\times 10^{-3}$)	G_n^ℓ ($\times 10^{-3}$)
Set 1 [a]	b_4	b_3	b_2	b_1	μ_4^ℓ	μ_3^ℓ	μ_2^ℓ	μ_1^ℓ	(USING f = 0.75)	EXPERIMENTAL
	9.65				3.5				0.093	0.093
	9.65	8.51			3.5	2.9			1.5	1.5
	9.65	8.51	7.62		3.5	2.9	2.0		17	18
	9.65	8.51	7.62	6.86	3.5	2.9	2.0	2.5	250	250
Manufacturer's estimate	9.65	8.51	7.62	6.86	(7.0)	(3.5)	(2.5)	(2.0)	(590)	---
Set 2 [b]	9.65	8.51	7.62	6.86	2.0	1.8	1.7	1.2	50	50
Manufacturer's estimate	9.65	8.51	7.62	6.86	(7.0)	(3.5)	(2.5)	(2.0)	(590)	---
Set 3 [c]	23.88	21.84	20.07	18.54					---	60
Manufacturer's estimate	23.88	21.84	20.07	18.54	(10)	(5.0)	(2.5)	(2.0)	(100)	---

(a) Mu Shield Corp. Malden, Mass. 02148 2.54 cm access hole, 0.051 cm wall thickness

(b) Perfection Mica Corp. Bensenville, Ill. 60106 2.54 cm access hole, 0.051 cm wall thickness

(c) Allegheny Ludlum Brackenridge, Pa. 15014 6.35 cm access hole one end
4.826 cm access hole other end
0.081 cm wall thickness

QUESTIONS AND ANSWERS

DR. VICTOR REINHARDT, NASA Goddard Space Flight Center:

I have a few questions. First of all, on your holes, I noticed that you had some little flanges. Did you experiment to determine if those flanges improved the shielding factor?

MR. WOLF:

Well, it turns out that we did experiment with that. On our initial design, we felt these flanges would make a difference. But we had a set of shields fabricated without the flanges, and it turned out that there was no difference.

DR. REINHARDT:

Okay. Another question: The variations in the μ. Did you find a lot of variation in the samples from the same manufacturer?

MR. WOLF:

We actually looked at three manufacturers. The μ values of two of the manufacturers were quite comparable once the shields were annealed to their best state. In one case, we had to have the shields reannealed a second time.

The third manufacturer, the μ value was about a factor of two lower than the other two, even though the original specifications were the same. I suspect that there would be a large variation in the μ that you could get from manufacturer to manufacturer.

DR. REINHARDT:

No, but within the same manufacturer, did you find reproducible μ's if you ordered the same set of shields?

MR. WOLF:

Yes. Once the shields were properly annealed, the permeability was quite constant.

DR. JACQUES VANIER, Laval University:

First, when you mentioned the part in 10^{14} that you required, what was the field fluctuation you assumed?

MR. WOLF:

Okay. I'm sorry; I should have mentioned that. We assumed a field variation, an external field variation, of ±1 gauss.

DR. VANIER:

Okay. What method of degaussing these shields did you use? Could you comment on that?

MR. WOLF:

We tried many. The best method was actually a twofold technique. We depermed the outer shields by placing them inside a ten-foot Helmholtz pair and put on an AC field of 30 gauss. And then after that, we took a wire and ran it through the inside and put on an AC field of about, again, 30 gauss, slowly decreasing the field. And we did it maybe two or three times, until the internal field was less than we could measure with our instruments. It was degaussed to about 10^{-6} gauss; a microgauss.

DR. VANIER:

Do you know the frequency of the degaussing?

MR. WOLF:

Yes. It was 60 cycles.

DR. GIOVANI BUSCA, Ebauches, Switzerland:

Did you find some problem in the joint of the shields? Normally, people say that the joint is the most critical part of the shields.

MR. WOLF:

Well, it turns out that we did not see any difference. The shields that we got for the SAO-VOG-11 shields were welded on one end with just a mechanical joint on the other. The fact that we got such good agreement with the profile for just taking into account the size of the holes indicated that there wasn't very much difference between the spot-welded joint, and just the mechanical joint.

ELIMINATION OF THE LIGHT SHIFT IN RUBIDIUM GAS CELL FREQUENCY STANDARDS USING PULSED OPTICAL PUMPING

Thomas C. English, Ernst Jechart and T. M. Kwon
Efratom Systems Corporation, Irvine, Ca.

ABSTRACT

It is well known that changes in the intensity of the light source in an optically pumped, rubidium, gas cell frequency standard can produce corresponding frequency shifts, with possible adverse effects on the long-term frequency stability. Since this so-called "light-shift" effect is due to the simultaneous presence of pumping light and interrogating microwave radiation, it can be eliminated, in principle, by alternately pulsing the pumping light and the microwave radiation so that there is no temporal overlap.

We have constructed a pulsed optical pumping apparatus with the intent of investigating the frequency stability in the absence of light shifts. Contrary to our original expectations, a small residual frequency shift due to changes in light intensity has been experimentally observed. Evidence is given which indicates that this is not a true light-shift effect. Preliminary measurements of the frequency stability of this apparatus, with this small residual "pseudo" light shift present, are presented. It is shown that this pseudo light shift can be eliminated by using a more homogeneous C-field. This is consistent with the idea that the pseudo light shift is due to inhomogeneity in the physics package ("position-shift" effect).

INTRODUCTION

At the present time, there is a real need for compact, lightweight, low power, highly stable atomic frequency standards. This is ex-

emplified by programs such as that for the NAVSTAR Global Positioning System, or GPS for short. This program calls for state-of-the-art size, weight and power reductions for atomic frequency standards, and a long-term stability requirement of one part in 10^{13} over periods of approximately 1 to 12 days for phases II and III of GPS {1}. Requirements such as these can be expected to be even more demanding in the future as the overall state-of-the-art of military technology moves forward.

Table 1 summarizes the state-of-the-art for small atomic frequency standards. The hydrogen and cesium devices have excellent stability over periods of weeks, months and even years. Rubidium is less satisfactory in this respect but has the advantage of much smaller size, weight and power consumption. The ideal frequency standard would combine the excellent long-term stability of hydrogen and cesium with the small size, weight and power consumption of rubidium. There are two possible methods of approach toward this goal. The first is to reduce even further the size, weight and power of hydrogen and cesium devices. While further small reductions might be possible, it is unlikely that these devices can be made to approach present day rubidiums in this respect. For example, even for cesium a factor of 9 in size and a factor of 7 in weight would be required. Moreover, relative to future rubidiums, factors of 18 and 12, respectively, would be required {2}. We are therefore led to the second method of approach, which is to improve the long-term stability of rubidium devices without significantly increasing size, weight, or power consumption. Table 2 shows that there is reason to believe that this can be done.

The best reported stability for rubidium is several parts in 10^{14} for an averaging time of about 7 hours. Two such measurements have been reported {3,4}, each on a different unit at different times and by different groups. In addition, a stability of parts in 10^{14} for τ = 6 hr to 12 days for one rubidium device has been reported at this conference (this device uses an Efratom physics package) {1}. This stability is on a par with cesium {5} and is only about a factor of 50 worse than the "best" reported stability for hydrogen {6}. For times longer than 7 hours, the stability of most rubidium devices worsens, most likely due to uncontrolled changes in device parameters. If the device parameters that are changing could be determined and controlled, better long-term stability would result.

FUNDAMENTAL PHYSICAL EFFECTS

Figure 1 lists those fundamental physical effects that can adversely affect the stability of rubidium devices. Let us consider these effects one at a time.

Light Shift

The insert at the lower left of Figure 1 shows the two hyperfine energy levels of rubidium which determine the rubidium resonant frequency. In general, when a rubidium atom is illuminated by the light used for optical pumping, these two energy levels are Stark shifted by the electric field of the light, thereby producing a change in the atom's resonant frequency {7}. This results in a frequency shift that is proportional to light intensity.

This effect can be eliminated by interrogating the atoms in the dark; that is, when the pumping light is absent.

Position-Shift Effect

The position-shift effect {8} is due to inhomogeneity in the physics package. The result of this inhomogeneity is to make the resonant frequencies of the rubidium atoms depend on their location within the cell, as indicated schematically at the lower right of Figure 1. This inhomogeneity is due mostly to C-field nonuniformity and also, to a lesser extent, to light shift nonuniformity. Since the optical pumping and the microwave interrogation are also nonuniform over the cell, the experimentally observed resonant frequency is a weighted average of the frequencies of the individual atoms in the cell. Now, if the intensity of pumping light were to change, then the weighted average would for example, shift so that those atoms in the center of the cell might be favored more than those at the ends. This would obviously result in an increase in the experimentally observed resonant frequency. We term this type of shift the "pseudo light shift" because it can mimic a true light shift.

Buffer Gas Shift

The buffer gas shift {9} occurs due to collisions between the rubidium atoms and the buffer gas atoms, and is dependent on buffer gas density and temperature, as well as on the buffer gas that is used.

Spectrum of Exciting Microwave Radiation

Frequency shifts due to the exciting microwave radiation are not usually a problem if care is taken to obtain a spectrally pure exciting frequency free of spurious and unwanted sidebands {10}.

Magnetic Field

Since the magnetic field sensitivity {10} for rubidium is only about 2 x larger than for cesium, magnetic field sensitivity for rubidium

is not significantly more of a problem than it is for cesium.

It is very difficult to give estimates of the possible frequency changes due to these effects because they are a very strong function of the individual device configuration and parameters. In spite of this, we can still say that of the 5 effects listed here, the first 3 are the most important.

The objective of the experiments that we are carrying out is to improve the long-term stability of small rubidium devices. Today we will describe some preliminary results that lead toward this goal. These results deal with the reduction and elimination of the light-shift and the position-shift effects, and have been obtained with the expenditure of less than one man-year of scientific effort.

METHOD FOR ELIMINATION OF LIGHT SHIFT

Figure 2 shows the method that we have used to eliminate the light shift. This method was first suggested by Arditi {11}. First, the light is pulsed on and the atoms are optically pumped. Then the light is turned off and the interrogating microwave radiation is turned on. The microwave radiation is then turned off and the pumping light is turned on again. This basic cycle is subsequently repeated many times. Since the atoms are interrogated in the dark, the light shift should be eliminated. This, of course, assumes that the atomic coherence does not persist from one cycle to the next {12}.

Before passing to the next slide, we note that in our experiments, the atomic transition is detected by optically monitoring the absorption of the pumping light. This is essentially the same detection method as that used in all conventional rubdiums.

Figure 3 shows a block diagram of the apparatus. The pulser alternately pulses the light and the microwave radiation at a 280 Hz rate, as shown in the previous slide, so that the atoms are interrogated in the dark. The remainder of the apparatus is a conventional frequency locked loop. The modulation frequency of 10 Hz is chosen to be about an order of magnitude smaller than the pulsing frequency so that the two signals can be separated by filtering before synchronous detection of the 10 Hz.

Figure 4 is a photograph of the physics package, which is a modified version of the physics package used in the Efratom, Model FRK rubidium frequency standard. The base of the rubidium lamp is at the right, and the magnetic shield that encloses the resonance cell and microwave canity is at the left. The entire unit is less than 4 inches long.

The philosophy adopted in this work was that everything possible should be done to retain the small size.

RESULTS

Light shift measurements have been made on this apparatus and the results are shown in Table 3. Measurements were made of the fractional frequency shift resulting from a 30 % change in light intensity. Two sets of measurements were made - - one with the apparatus operated CW, and the other with it pulsed.

In the case of CW operation, we expect a large light shift. This was observed for each of two different rubidium lamps - - lamp A and lamp B. These two lamps differ in the ratios of their rubidium isotopes. Lamp A produces a positive light shift, and lamp B a negative light shift. This is in agreement with the theory of the light shift in rubidium 87 as worked out by Mathur, Tang and Happer {7}.

When the apparatus is operated in the pulsed mode we expect to see no light shift. Yet there is a change in frequency with light intensity. This change is about a factor of 10 smaller than the CW light shift, and we can tell that it is not a true light shift because it does not change sign in going from lamp A to lamp B. Other tests, which we will not describe here, also confirm this to be the case. For these reasons we dub this effect the "pseudo-light-shift effect." We will have more to say on this later.

Table 4 shows the results of some preliminary frequency stability measurements that have been made on our pulsed optical pumping apparatus. For pulsed pumping, the short-term stability is expected to be degraded somewhat by noise introduced in the pulsing process. The short-term stability for pulsed pumping has been measured for averaging times from 1 to 100 seconds and found to improve as $1/\sqrt{\tau}$ (footnote A in Table 4). This shows that we are dealing with white frequency modulation noise, as is usually the case for passive rubidium devices. The value of $\sigma_y(\tau)$ for 100 sec is given in column 2 and can be compared with that for our small commercial rubidiums. The result for pulsed pumping lies between the spec for our two commercial models and is better than that of an HP 5062C cesium. The short-term stability of the pulsed pumping apparatus is therefore quite good, even in this preliminary stage, and can almost certainly be improved further.

The long-term stability was also measured for a 24-hour averaging time and a preliminary value of approximately 5 parts in 10^{12} was obtained. This is not yet as good as our commercial units.

After these stability measurements were made it was discovered that there were several device parameters that were not under tight control. These included significant second-harmonic contamination of the 10 Hz modulation, and frequency changes due to changes in barometric pressure. All of these parameter changes can be expected to produce frequency changes of parts in 10^{12}, which is of the order of the observed instability over 24 hours. The pseudo-light-shift effect is not negligible at this level either, and it may be a contributor to the observed instability.

All of these parameters can be easily controlled except for the pseudo light shift. However, it was suspected that the pseudo-light-shift effect might actually be a manifestation of the position-shift effect, as mentioned earlier. To test this hypothesis, a new C-field was constructed for our physics package that greatly improved the homogeneity of the static magnetic field and which should therefore greatly diminish the position-shift effect.

Table 5 shows the result of using this new C-field. The first line of this table is a repeat of the data shown in Table 3 for the old C-field. The second line shows what happened when some small steel parts on the outside of the microwave cavity were removed. This improved the C-field homogeneity and also reduced the pseudo light shift by about 30 %. Finally, the last line of the table shows the results for the new C-field. The pseudo light shift is now undetectable, of the order of parts in 10^{12} or less for a 30 % change in light intensity.

To summarize, the residual light shift has now been reduced to an undetectable amount by using the method of pulsed optical pumping in conjunction with a homogeneous C-field. It is likely that the new homogeneous C-field will also have other beneficial effects, such as reduced sensitivity to changes in microwave power, but this has not yet been verified experimentally.

Our plans for the immediate future are to beat down the known sources of frequency instability to the level of parts in 10^{13} or below, and then to take additional long-term stability data. It is expected that this will lead to an improved long-term stability compared to the present value of about 5 parts in 10^{12} for $\tau = 24$ hours which was taken before the pseudo light shift was eliminated. Stability data over longer periods of time will also be taken to see if there is an improvement there.

Table 6 compares our preliminary results with those obtained by other investigators, namely, Arditi and Carver, who were the first to use the method of pulsed optical pumping for elimination of the light shift, and a Russian group that has done several man-years of work in

this area.

In their experiments, Arditi and Carver used a high sensitivity microwave receiver to detect the rubidium resonance. Because of the complexity of the electronics this method is not suitable for use in a practical device. The Russians used an optical detection method, that has the disadvantage of requiring 2 rubidium lamps. We also use an optical detection method but only one rubidium lamp is required. In addition, we have used a single rubidium cell that combines the filtering and resonance functions, thereby eliminating the need for a separate filter cell. For these reasons our physics package is extremely small, which is desirable in a practical device. In fact, this is the same physics package that is used in the Efratom small commercial rubidium frequency standards.

Arditi and Carver saw no light shift at the level of a part in 10^{10}. On the other hand, the Russians did observe true light shifts (due to persistence of the atomic coherence) but were able to minimize them by proper choice of operating conditions. In our experiments the light shift is undetectable so that if it exists at all, it is of order parts in 10^{12} or less for a 30 % change in light intensity.

As regards stability, Arditi and Carver made some short-term measurements at the level of about 1×10^{-10}. The Russians have not reported any stability measurements for reasons unknown to us. As already mentioned, we have measured the frequency stability for our apparatus and found it to be of the order of parts in 10^{12}. However, it should be emphasized that these measurements are preliminary and were made prior to elimination of the pseudo light shift.

NOTE ADDED IN PROOF

We have just learned that results similar to those reported here have also been obtained by J. Ernvein-Pecquenard and L. Malnar, "Horologe atomique à pompage optique séquentiel," C. R. Acad. Sc. Paris 268(B), 817 (1969). They reported no light shift (detection limit of $1 - 2 \times 10^{-11}$) for a 25 % change in light intensity using pulsed optical pumping with optical detection; no frequency stability results are given in their paper, however. Also, no pseudo light shift was observed in their experiments, presumably because they used a conventional, laboratory-type, optical pumping apparatus (very large physics package and very homogeneous C-field); for practical devices utilizing small physics packages, it is also necessary to consider the pseudo light shift effect (position-shift effect), as has been done in the present work. Thus, the results of the two investigations are in agreement, to the extent that the experimental setups were similiar. We would like to thank Prof. M. Têtu for bringing the paper of Ernvein-Pecquenard and Malnar to our attention.

PROGNOSIS

At the present time, the ultimate frequency stability attainable using the pulsed optical pumping method is not known. Possible limitations could be due to changes in buffer gas pressure, if they occur, and also to changes in the pulsing parameters, such as the durations of the light and microwave pulses. It is known that frequency shifts do occur due to changes in the pulsing parameters. We have investigated this phenomenon using the old C-field and have estimated that it certainly is important at the level of parts in 10^{13}.

Future efforts will concentrate on understanding and reducing frequency sensitivity to changes in pulsing parameters. Effort will also be devoted to devising a method for studying frequency shifts due to possible small changes in buffer gas pressure.

ACKNOWLEDGEMENTS

We would like to thank Werner Weidemann, Engineering Manager of Efratom Systems Corporation, for many helpful discussions, and for kindly making available both equipment and facilities, as needed. We also acknowledge the able technical assistance of Jeff Hayner, Henry Holtermann and John Hall.

REFERENCES

1. F. K. Koide, D. J. Dederich and L. H. Baker, "Rubidium Frequency Standard Test Program for NAVSTAR GPS," Paper 9 presented at Session III of this conference (10th PTTI, 1978).

2. T. C. English, presentation on rubidium frequency standards as part of the Discussion Form on Atomic Frequency Standards, Session II of this conference (10th PTTI, 1978).

3. H. Hellwig and A. E. Wainwright, "A Portable Rubidium Clock for Precision Time Transport," 7th PTTI, Dec 1975, p. 143.

4. Unpublished data, Rockwell International, Anaheim, California, 1977.

5. D. B. Percival, "Statistical Properties of High Performance Cesium Standards," 5th PTTI, Dec. 1973, p. 239.

6. M. W. Levine, R. F. Vessot and E. M. Mattison, "Performance Evaluation of the SAO VLG-11 Atomic Hydrogen Masers," Proc. 32nd Annual Symp. on Freq. Control, 31 May-2 June, 1978, p. 477.

7. B. S. Mathur, H. Tang and W. Happer, "Light Shifts in the Alkali Atoms," Phys. Rev. 171, 11 (1968).

8. A. Risley and G. Busca, "Effect of Line Inhomogeneity on the Frequency of Passive Rb^{87} Frequency Standards," Proc. 32nd Annual Symp. on Freq. Control, 31 May-2 June, 1978, p. 506.

9. W. Happer, "Optical Pumping," Rev. Mod. Phys. 44, 169 (1972), espec. pp. 221-24.

10. R. C. Mockler, "Atomic Beam Frequency Standards," Adv. Electronics and Electron Phys. 15, 1 (1961), espec. pp.51. Also, reprinted in NBS Special Publication, Volume 5, Precision Measurement and Calibration, Frequency and Time, 1972, p. 1.

11. It appears that the idea of pulsed optical pumping was originally conceived by Alley {C. O. Alley, "Coherent Pulse Techniques in the Optical Detection of the $0 \leftrightarrow 0$ Ground State Hyperfine Resonance in Rubidium 87," in Quantum Electronics (C. H. Townes, Editor, Columbia U. Press, N. Y., 1960), p. 146} . The application for the elimination of the light shift appears to have been due to Arditi {M. Arditi, "High Stability Gas Cell Frequency Standard," U. S. Patent 3, 234, 483, Feb. 8, 1966} .

12. E. I. Alexseyev, Y. N. Bazarov and G. I. Telegin, "Light Shifts in a Quantum Frequency Standard with Pulsed Optical Pumping and Optical Indication of the Ramsey Structure of the $0 \leftrightarrow 0$ Transition Line of Rb^{87} Atoms," Radio Eng. and Electronic Phys. 20, 73 (1975)

TABLE 1

SOME CHARACTERISTICS OF SMALL STATE-OF-THE-ART

ATOMIC FREQUENCY STANDARDS

DEVICE	LONG-TERM STABILITY	SIZE (LITERS)	WEIGHT (LBS)	POWER (W)	APPROX. COST (K$)
SPACECRAFT H-MASER	$< 1 \times 10^{-14}$/10 D[A]	20-50[A]	50-90[A]	55[A]	- - -
SMALL COMMERCIAL CESIUM	PARTS IN 10^{12}/YR[B]	9	22	24	18
SMALL COMMERCIAL RUBIDIUM	$< 1 \times 10^{-11}$/MO[B]	1	3	13	6

[A]PROJECTED

[B]TYPICAL MANUFACTURER's SPECIFICATION

TABLE 2

"BEST" REPORTED FREQUENCY STABILITIES FOR

ATOMIC FREQUENCY STANDARDS

DEVICE	STABILITY, $\sigma_Y(\tau)$	AVERAGING TIME, τ	REF
SAO H-MASER	6×10^{-16}	11 HR	1
COMMERCIAL CESIUM (HIGH PERFORM)	2×10^{-14}	5 D	2
SMALL COMMERCIAL RUBIDIUM	$3\text{-}4 \times 10^{-14}$	7 HR	3,4
GPS SPACE CRAFT RUBIDIUM (S/N 2)	$\leq 8 \times 10^{-14}$	6 HR TO 12 D	5

REFERENCES:
1. LEVINE ET AL., 32 ND ANNUAL SYMPOS. ON FREQ. CONTROL, 1978, P. 477.
2. PERCIVAL, 5 TH PTTI, 1973, P. 239.
3. HELLWIG AND WAINWRIGHT, 7 TH PTTI, 1975, P. 143.
4. UNPUBLISHED DATA, ROCKWELL INTERNATIONAL, ANAHEIM, CALIFORNIA 1977.
5. KOIDE ET AL., PAPER 9, SESSION III, THIS CONFERENCE (10TH PTTI, 1978).

TABLE 3

PSEUDO LIGHT SHIFT EFFECT FOR PULSED OPTICAL PUMPING

RB LAMP	FRACTIONAL FREQUENCY SHIFT DUE TO LIGHT INTENSITY CHANGE $(Y)_{I_o} - (Y)_{0.7 I_o}$	
	CW	PULSED
A	$+1.4 \times 10^{-9}$	$+1.5 \times 10^{-10}$
B	-1.6×10^{-9}	$+2.4 \times 10^{-10}$

TABLE 4

PULSED OPTICAL PUMPING

PRELIMINARY STABILITY RESULTS

DEVICE	σ_Y (100 SEC)	σ_Y (24 HR)
PULSED PUMPING (PRESENT WORK)[A]	2×10^{-12}	$\sim 5 \times 10^{-12}$
EFRATOM FRK-H RUBIDIUM[B]	1×10^{-12}	$(< 1 \times 10^{-12})$[C]
FRK-L RUBIDIUM[B]	3×10^{-12}	- - -
HEWLETT-PACKARD 5062C CESIUM[B]	7×10^{-12}	$(< 1 \times 10^{-12})$[C]

[A] $\sigma_Y = 2 \times 10^{-11}\ \tau^{-1/2}$, $1 \le \tau \le$ 100 SEC

[B] MANUFACTURERS's SPECIFICATION

[C] UPPER LIMIT

TABLE 5

LIGHT-SHIFT MEASUREMENTS FOR DIFFERENT C-FIELD CONFIGURATIONS (LAMP A)

C-FIELD	FERROMAGNETICS ON CAVITY	FRACTIONAL FREQUENCY SHIFT DUE TO LIGHT INTENSITY CHANGE $(Y)_{I_o} - (Y)_{0.7I_o}$	
		CW	PULSED
OLD	YES	$+1.4 \times 10^{-9}$	$+1.5 \times 10^{-10}$
OLD	NO	$+1.3 \times 10^{-9}$	$+1.0 \times 10^{-10}$
NEW	NO	$+1.0 \times 10^{-9}$	$< 1 \times 10^{-11}$

TABLE 6

SUMMARY OF PULSED OPTICAL PUMPING RESULTS

STUDY	DETECTION SCHEME	LIGHT SHIFT	STABILITY DATA
ARDITI & CARVER[A] (1964)	MICROWAVE SUPERHET	NO ($< 1 \times 10^{-10}$)	YES $\sim 1 \times 10^{-10}$
ALEXSEYEV *ET AL*[B] (1974)	OPTICAL (TWO RB LAMPS)	YES	NO
PRESENT WORK	OPTICAL (SMALL PHYSICS PACKAGE)	NO ($< 1 \times 10^{-11}$)	YES (PRELIMINARY) TO PARTS IN 10^{12}

[A] IEEE Trans. Instr. Meas., June - Sept., 1964, p. 146.

[B] Radio Eng. Elec. Phys. 20, 73 (1975).

FIGURE 1

FUNDAMENTAL PHYSICAL EFFECTS THAT CAN ADVERSELY AFFECT THE LONG-TERM STABILITY OF PASSIVE RUBIDIUM DEVICES

- LIGHT SHIFT($\sim 3 \times 10^{-11}$ FOR $\Delta I_{LIGHT}/I_{LIGHT} = 1\ \%$)
- POSITION-SHIFT EFFECT (UP TO PARTS IN 10^9)
- BUFFER GAS SHIFTS ($\sim 1 \times 10^{-10}$/MILLITORR)
- SPECTRUM OF EXCITING MICROWAVE RADIATION (UP TO PARTS IN 10^9)
- MAGNETIC FIELD (RB SENSITIVITY = 1.8 x CS SENSITIVITY)

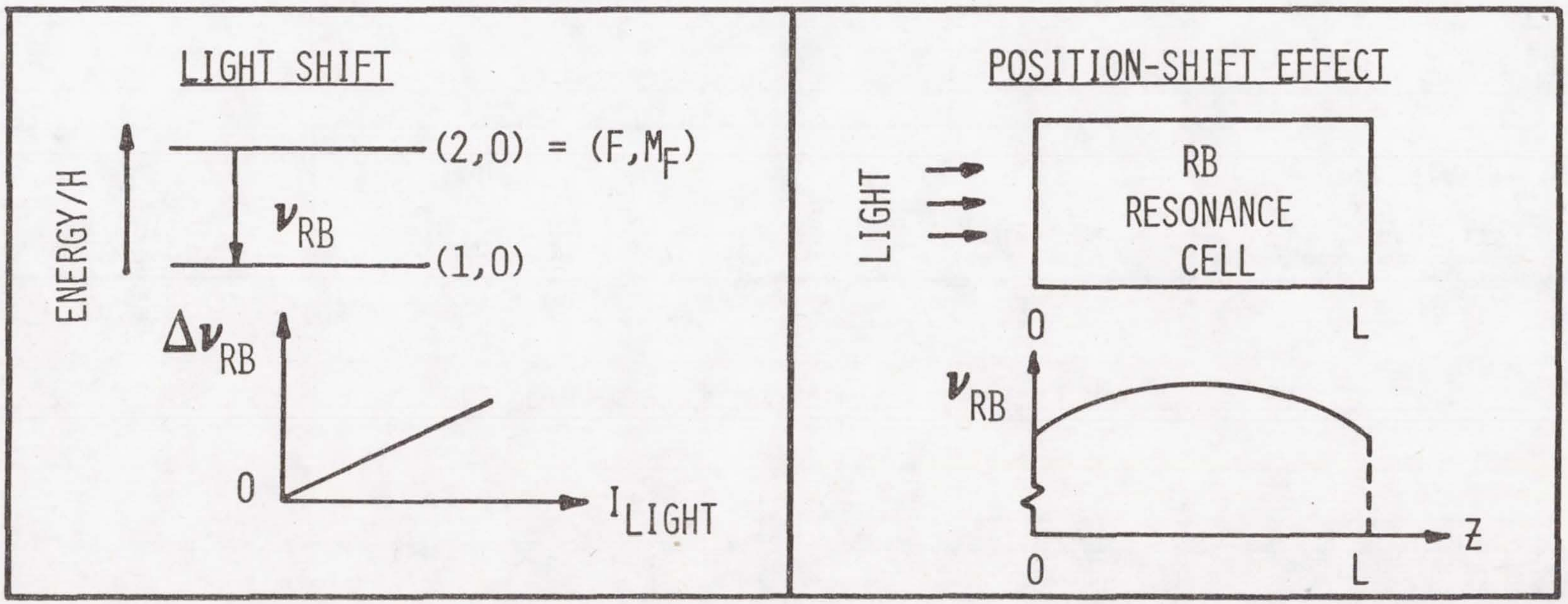
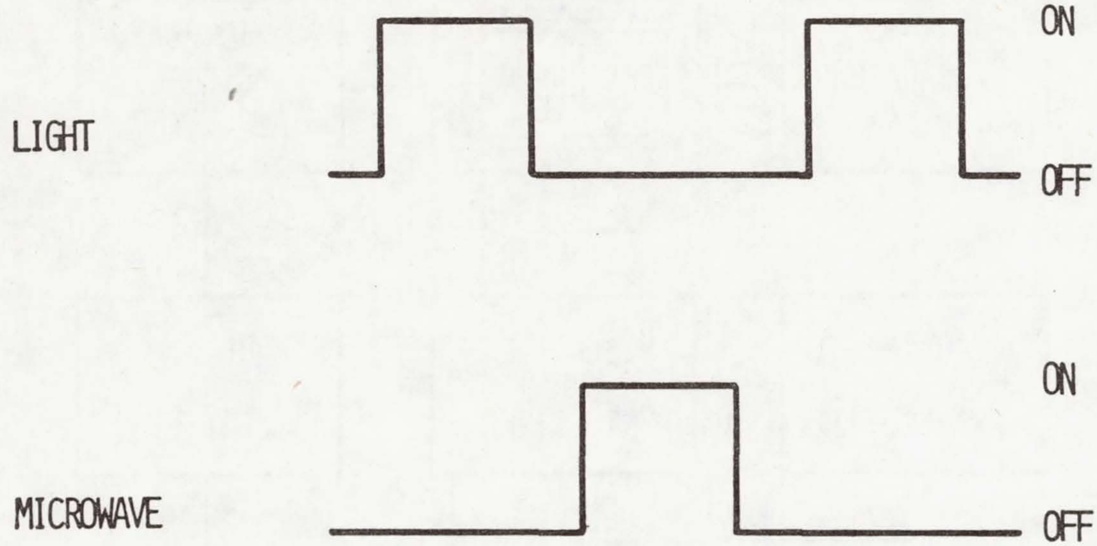


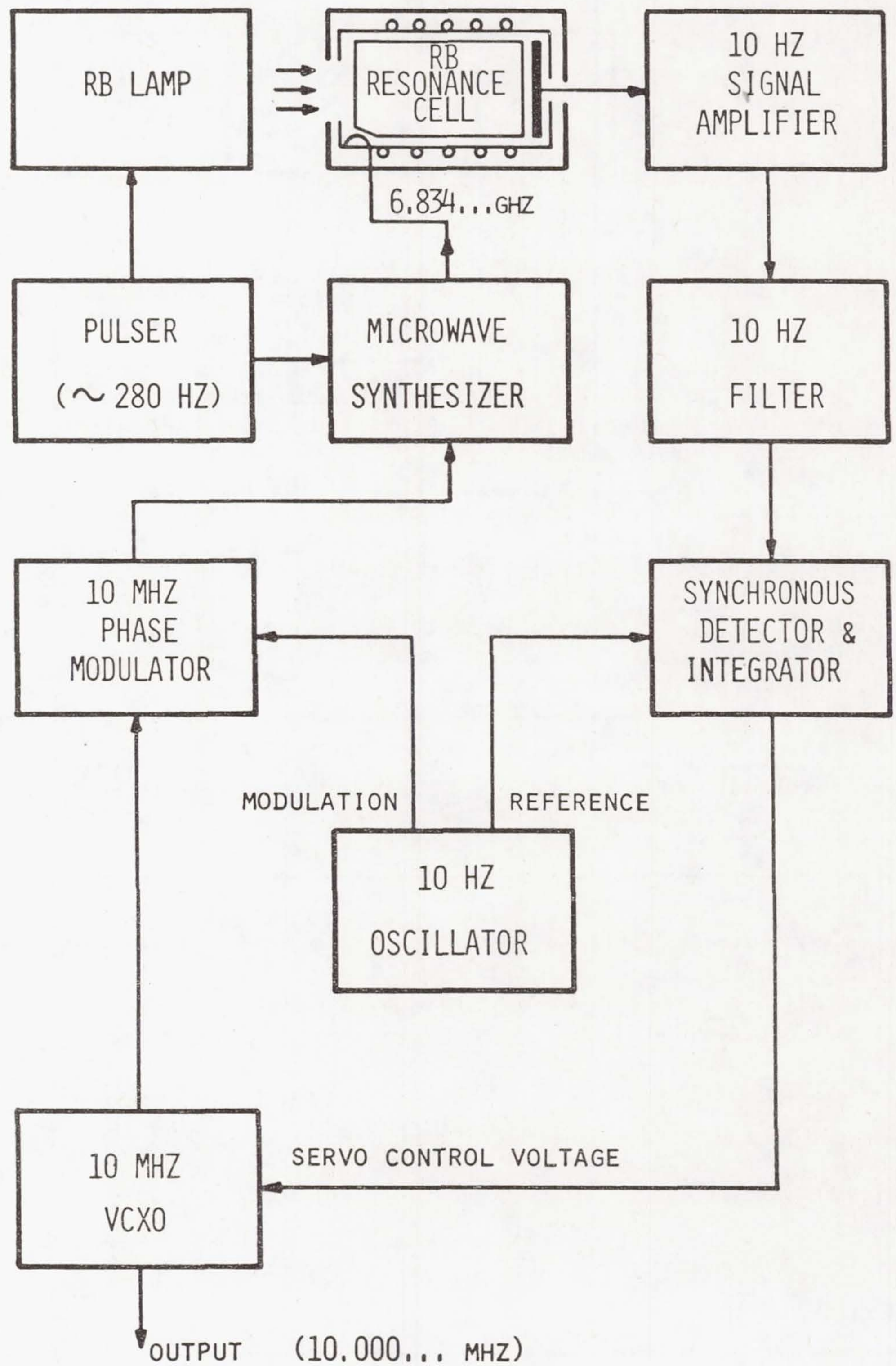
FIGURE 2
PULSING SCHEME



PULSE PARAMETER	TYPICAL VALUE (MSEC)
LIGHT DURATION	2.0
DARK TIME	1.6
MICROWAVE DURATION	1.2
REPETITION RATE	280 HZ

FIGURE 3

BLOCK DIAGRAM OF APPARATUS



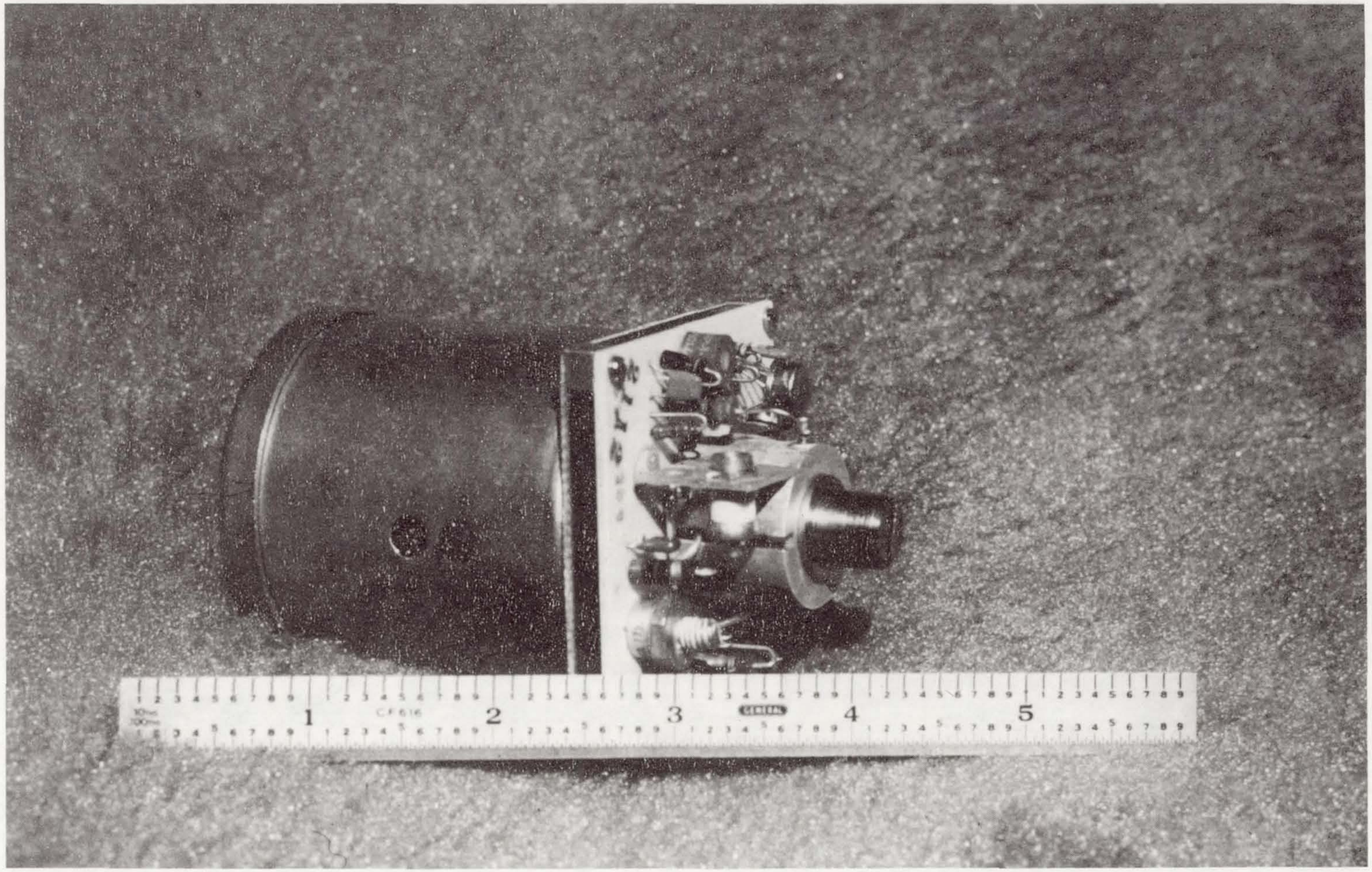


FIGURE 4. PHOTOGRAPH OF PHYSICS PACKAGE

QUESTIONS AND ANSWERS

DR. GIOVANI BUSCA, Ebauches, Switzerland:

I see you are using 10 hertz frequency for modulation. So you use phase sinusoidal modulation? What was the linewidth?

DR. ENGLISH:

Well, it is variable, and it depends upon how you choose the pulse parameters. Now if you run the units CW, it is on the order of a kilohertz. And if you really try to get the narrowest linewidth you can, where you are limited by the relaxation times for spin exchange, then you can actually get the linewidth down to a couple of hundred hertz. In these experiments, though, if you choose the narrowest linewidth, then your signal really goes to pot. So you work at some intermediate range. And we have been using roughly 700 hertz.

DR. BUSCA:

I see. I am asking myself if some of the results could be improved just by increasing the modulation frequency.

DR. ENGLISH:

It is possible. I wouldn't rule it out.

DR. BUSCA:

You are very far from the optimum value, which is. . . .

DR. ENGLISH:

Yes, you are absolutely right. It affects the signal-to-noise ratio, and it is possible if you went to a higher modulation frequency, you could increase your signal and therefore the short-term stability.

DR. BUSCA:

Okay. I have a second question. It seems that you use data on light shift when you use the Lamp in continuous CW and you extrapolate this data to a case in which you use the excitation pulse. We have done that before, and we see that the spectrum could change very much from pulsed to continuous excitation. So you must take a little bit of care in extrapolating from continuous data to pulsed data.

DR. ENGLISH: Yes.

DR. BUSCA:

The change in the sign, for example, from pulsed to continuous excitation would be just something like that. That means changing the spectrum of the light.

DR. ENGLISH:

Well, the Russian work seems to indicate--and they worked out a rather complicated theory for the whole thing--that you should have more or less a correspondence between the CW value and the pulsed value, not necessarily one to one, but at least a sign change is a strong indication that it is not a true light shift. In addition, there were the other factors which I don't have time to go into, which have to do with how the effect varies with the duration of the various pulse parameters.

I think the most convincing evidence that it is what I call a pseudo light shift is that it did disappear when all we did was improve the C-field homogeneity.

DR. BUSCA: Okay.

DR. GERNOT M. R. WINKLER, U.S. Naval Observatory:

I have a dim recollection that the scheme of using pulsed pumping originated in Princeton, 1958 or 1959, by Carroll Alley.

DR. ENGLISH:

It is quite possible. I know Arditti has a patent on this, and that is the basis for my saying that I believe he originated it. Now Arditti and Carver wrote a paper together on it. It is the first one I am aware of in regular journals. But you could be right. It could have been an idea that was sort of floating around Dicke's lab and they picked it up. I don't really know. I have never talked to him about it.

DR. HELMUT HELLWIG, National Bureau of Standards:

I only have one comment. Many things we are now working on are not really new ideas if you analyze them. What is a new idea, really, but the rather childish curiosity which makes us ask the same questions today that were asked five years, ten years, twenty years ago? And things have changed sufficiently that you might get answers you did not dream about.

DR. ENGLISH:

Could I make one comment? One of the reasons for doing this is that the light shift is a very messy effect. And so if you really want to study a device and vary the parameters in it to study the other effects, it is really desirable to get rid of the light shift. And

that is one of the major motivations behind the experiment--to get rid of this really messy effect. And then you are free to vary the other parameters.

Now you can devise rubidium devices that run CW where they have very little light shift. But you have to restrict yourself to a very specific operating point. And as soon as you mess with the parameters of the device, then you start changing the light shift as well as everything else. And the problem is to try and isolate the variables in such a way that when you change parameters in the device, you are not changing all of these different effects at once in ways that you can't separate.

PERFORMANCE OF PREPRODUCTION MODEL CESIUM BEAM FREQUENCY STANDARDS FOR SPACECRAFT APPLICATIONS

Martin W. Levine
Frequency & Time Systems, Inc.
Danvers, Massachusetts

ABSTRACT

The first of a series of Preproduction Model (PPM) Cesium Beam Frequency Standards for spaceflight application on Navigation Development Satellites has been designed and fabricated and preliminary testing has been completed. The PPM is identical in form, fit and function, to the production cesium standards for the GPS NAVSTAR System.

The PPM evolved from an earlier Prototype Model launched aboard NTS-2 (June, 1977) and the Engineering Development Model (EDM) to be launched aboard NTS satellites during 1979. A number of design innovations, including a hybrid analog/digital integrator and the replacement of analog filters and phase detectors by clocked digital sampling techniques will be discussed.

Thermal and thermal-vacuum PPM testing has been concluded and test data will be presented. Stability data for 10 to 10^4 seconds averaging interval, measured under laboratory conditions, will also be shown.

INTRODUCTION

The Preproduction Model (PPM) Cesium Beam Frequency Standards are the most recent version of a series of atomic frequency standards specifically designed for spacecraft applications in the Global Positioning Satellites (GPS). The first atomic frequency standards to be successfully operated on an orbiting satellite were the rubidium devices flown by the U.S. Naval Research Laboratory aboard NTS-1 in 1974. Two Prototype Model Cesium Standards were launched on NTS-2 in June, 1977.[1] Figure 1 is a photograph of the prototype units. The next stage in the evolution of the GPS cesium standards was the Engineering Development Model (EDM).[2] EDM Number 2 is currently installed and awaiting launch on the U.S. Air Force NDS-4 payload and EDM Number 3 is in test on the NDS-5 satellite. A photograph of the EDM Standard is shown in Figure 2.

The PPM is similar in size, weight and outward appearance to the EDM; internally there are a number of electrical and mechanical refinements. The PPM is intended to be representative, in all physical and performance aspects to the production model phase III cesium standards for GPS satellites. Figure 3 is a photograph of the qualification model PPM in the initial test phase.

DESCRIPTION

Two fundamental requirements for the GPS frequency standard are that the standard operate in space for at least five years without adjustment and that it be capable of operating in the specified radiation environment. These two requirements strongly influenced the design of the PPM, particularly in the servo loop and integrator.

An overall block diagram of the PPM cesium beam standard is shown in Figure 4. The basic building blocks are common to all cesium standards; a high-quality voltage-controlled crystal oscillator (VCXO), a phase-modulated frequency multiplier and a digital frequency synthesizer to generate the 9.19 GHz cesium hyperfine transition frequency, a cesium beam tube, a low-noise modulation signal amplifier and filter, a phase detector, an integrator, and power supplies and controllers. As indicated in the block diagram, these blocks can be loosely combined into four functional subsystems, servo/integrator, cesium-beam, r.f. and power supply.

Servo/Integrator Subsystem

The servo/integrator subsystem is shown in block diagram form in Figure 5. The modulated signal at the output of the beam tube is amplified and filtered by the first stage. The second stage is a commutative filter which samples the amplified beam signal at twice the modulation rate. The sample-and-hold filter is followed by a synchronous detector and the hybrid analog/digital integrator.

The integrator in a cesium beam standard is a problem in any application. Given the five-year life and the radiation hardening requirements of GPS, the problem is particularly severe. The loop gain and loop time constant dictate an integrator time constant, however, is limited by the bias current of the amplifier and the leakage resistance of the integrator capacitor and the circuit board. The hybrid integrator, shown in the block diagram of Figure 5, uses a combination of analog and digital techniques to circumvent these problems. The rate-of-change of the output voltage of an integrator can be increased by increasing the R - C time constant or, more simply, by attenuating the output by means of a resistance divider. The difficulty with the second approach is that the maximum output voltage

that can be obtained is reduced by the attenuation factor. The attenuation can be tolerated if the analog integrator is augmented by a digital-to-analog converter which is incremented (or decremented) by one count whenever the analog integrator reaches its upper (or lower) limit, the digital-to-analog converter is implemented in the PPM cesium standard by using latching relays as the digital switches. The relay contacts have no offset voltages or series resistances as do semiconductor switches, but more importantly, they provide a non-volatile memory which retains the last oscillator control voltage setting in the event of a radiation-induced transient. Each count of the D-to-A converter corresponds to approximately 10 mV, a $\Delta f/f$ of approximately 1×10^{-10} for the VCXO. Interpolation between discrete converter steps is provided by attenuating the ± 10 mV output range by 1024, to about ± 10 μV, and adding this voltage to the output of the D-to-A converter.

Cesium Beam Tube Subsystem

The cesium beam tube subsystem consists of the cesium beam tube, the C-field regulator, and the cesium oven controller. The beam tube is the FTS-1A; used in the Prototype Model and EDM standards as well as the 4000 series manufactured by Frequency & Time Systems.

The C-field current regulator uses a latching-relay-controlled D-to-A converter similar to that used in the digital integrator. The latching relays provide immunity to radiation induced transients while allowing ground-commanded adjustments to the field current. The resolution of the C-field is one part in 1024 which corresponds to a change in the normalized standard frequency of approximately 4×10^{-13}.

The cesium oven is maintained at a constant temperature by a variable-duty-cycle switching regulator. The cesium oven temperature is sensed by a thermistor on the cesium oven located within the cesium tube;

the thermistor forms one leg of a resistance bridge which is balanced at the oven setpoint temperature.

R.F. Subsystem

The 10.23 MHz signal for the satellite GPS systems is generated by doubling the 5.115 output frequency from a modified FTS Model 1000 precision voltage-controlled crystal oscillator.[3] The doubler output is carefully filtered to suppress the 5.115 MHz sub-harmonic by at least 100 dB.

The 36th harmonic of 5.115 MHz, at 184.14 MHz is generated by the low-order frequency multiplier. The input stage of the low-order multiplier is square-wave phase (frequency-impulse) modulated at approximately 450 Hz, the cesium beam tube line-width. The output stage of the multiplier in turn is phase modulated by the 14.36...MHz output of the digital frequency synthesizer. The 184.14 MHz carrier, with its complex phase modulation spectrum, is fed to the X50 high-order multiplier; the first sideband, 9192.631770 MHz phase-modulated at 450 Hz, is selected by an output filter to excite the Ramsey cavity within the cesium beam tube.

Power Supply Subsystem

The PPM power supplies provide +15V, -15V, and +5V to operate the servo, and r.f. electronic circuitry. A separate +24V supply is used for the VCXO and two high-voltage supplies to power the cesium tube ion pump and the electron-multiplier.

All of the PPM supplies are preregulated by a high efficiency switching regulator followed by a series-pass regulator. The output of the preregulator feeds two transformer-coupled inverter stages which supply all of the required voltages.

The preregulator switching regulator and both inverters are clocked from a stable oscillator at approximately 37 KHz. The clock frequency is

selected to minimize power supply ripple at the Zeeman frequency and the harmonics of the Zeeman frequency.

The normal power consumption is approximately 24 watts after a one-hour warmup period. The maximum power consumption of the PPM is approximately 44 watts at startup.

Mechanical Construction

The mechanical design of the PPM is very similar to that of the EDM. The PPM package is 15 inches (38.1 cm) long, 7.6 inches (19.4 cm) high, and 5.1 inches (13 cm) wide and weighs 25.5 lbs. (11.6 kg).

The PPM packaging concept differs in only relatively minor respects from the EDM; the PPM circuit boards are connectorized, rather than hard-wired, and the cesium tube in the PPM has a soft mount. Although the FTS-1A cesium beam tube used in the PPM survived the 23 g qualification-level vibration test when hard-mounted to the shake table, the chassis structure resonances increase the vibration levels imposed on the installed tube to the point that it seemed prudent to shock mount the tube. The shock mounting consists of relatively thin elastomer strips sandwiched between stainless steel clamping bands.

QUALIFICATION REQUIREMENTS

The PPM Cesium Beam Frequency Standard must survive the vigors of the launch environment and then perform within specified limits for a minimum of five years in space.

Environmental Requirements

Qualification testing of the PPM, scheduled to begin on 27 November 1978 includes vibration, shock, EMI, thermal vacuum cycling, phase noise, and temperature stability tests. The qualification vibration levels

for the PPM are shown in Figure 6. Three minutes of random vibration in each of three mutually perpendicular axis are required.

The qualification levels for a three-axis simulated and pyrotechnic shock test are shown in Figure 7. The thermal profile for the long-term thermal cycling test is shown in Figure 8. Twenty-four full cycles at atmospheric pressure and eight cycles in vacuum are required for qualification.

Performance Requirements

The PPM performance limits with respect to frequency stability, frequency versus temperature spurious signal levels and phase noise must be maintained over a baseplate temperature range of 20 to 45°C.

Figure 9 is a plot of the frequency stability requirements as a function of averaging time, from one second to 10^5 seconds.

Figure 10 shows the acceptable spurious signal levels as a function of frequency offset from the carrier.

Figure 11 is a plot of the phase noise requirement as a function of frequency offset from the carrier.

The maximum allowable temperature coefficient of frequency for the PPM is 5×10^{-14} per degree C averaged over any 20 degree interval within the 20 to 45°C operating range.

PRELIMINARY PERFORMANCE DATA

The temperature coefficient of frequency of the PPM qualification model has been measured in vacuum with the results shown in Figure 12. The coefficient is approximately $1.5 \times 10^{-14}/^{\circ}C$ averaged over the +15 to +45°C temperature range, significantly better than the 5×10^{-14} per degree C contractual requirement. The frequency offset at 25°C, in air at atmospheric pressure, is shown on the figure to indicate the

measured change in frequency from air to vacuum at constant baseplate temperature.

The measured frequency stability, expressed as the Allan variance for averaging intervals of 10 to 10^4 seconds is shown in Figure 13. The broken line on the figure represents the contractual requirement; again the PPM performance exceeds the minimum requirements by a substantial margin. The phase noise as a function of offset frequency from the carrier has not been measured for the PPM. However, since the phase noise characteristics at offset frequencies greater than one Hz are determined solely by the crystal oscillator, the EDM data [3] is representative of the phase noise performance expected for the PPM.

The author wishes to acknowledge the support of the Office of Naval Research under contract N00014-74-C-0061, for portions of this work.

References

1. J. White, F. Danzy, S. Falvey, A. Frank, J. Marshall, "NTS-2 Cesium Beam Frequency Standard for GPS", Proceedings of the Eigth Annual PTTI Applications and Planning Meeting, December, 1976.

2. T. Gregory, "New Cesium Beam Frequency Standards for Flight and Ground Applications", Proceedings of the 31st Annual Symposium on Frequency Control, June, 1977.

3. D. Emmons, "A New Rugged Low-Noise High Precision Oscillator", Proceedings of the Eigth Annual PTTI Applications and Planning Meeting, December 1976.

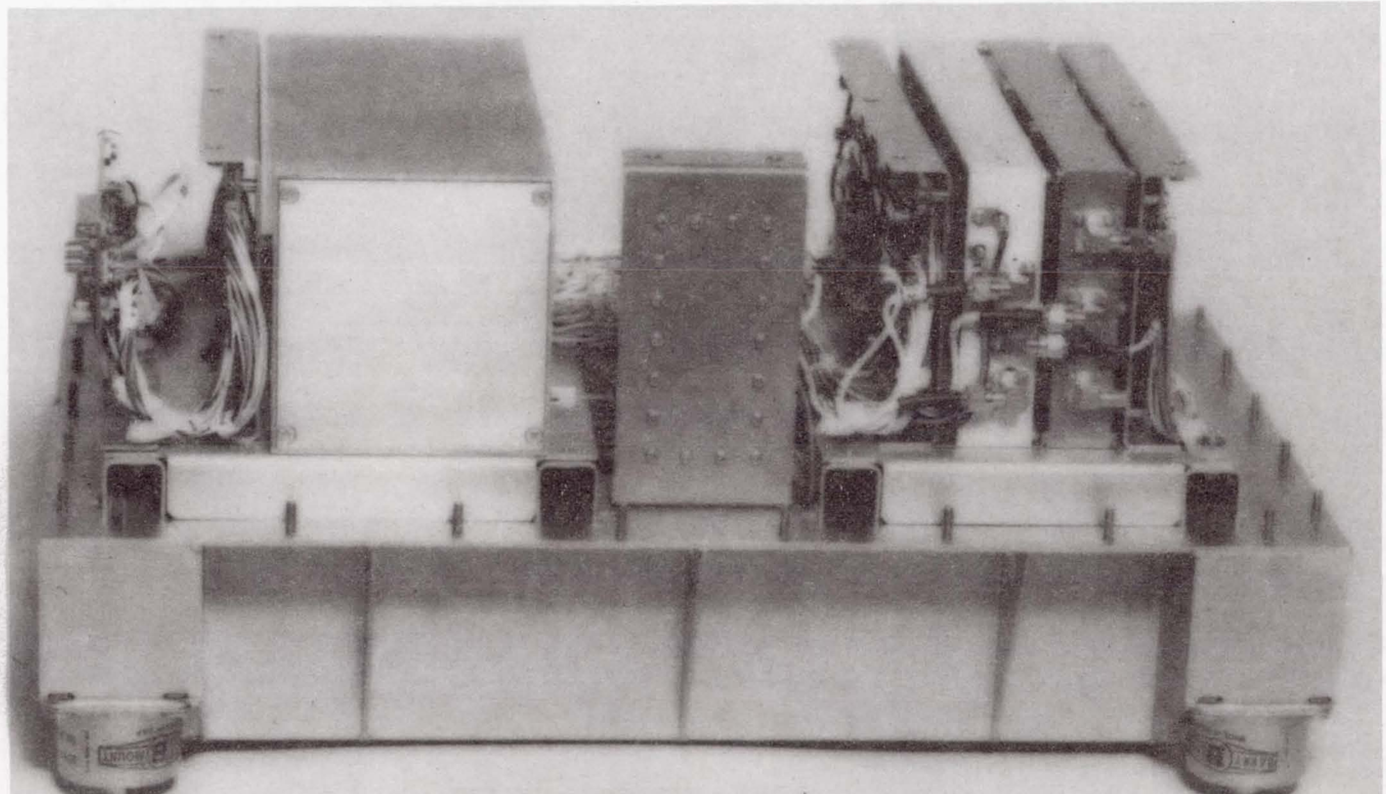


Figure 1
PHOTOGRAPH OF
PROTOTYPE MODEL

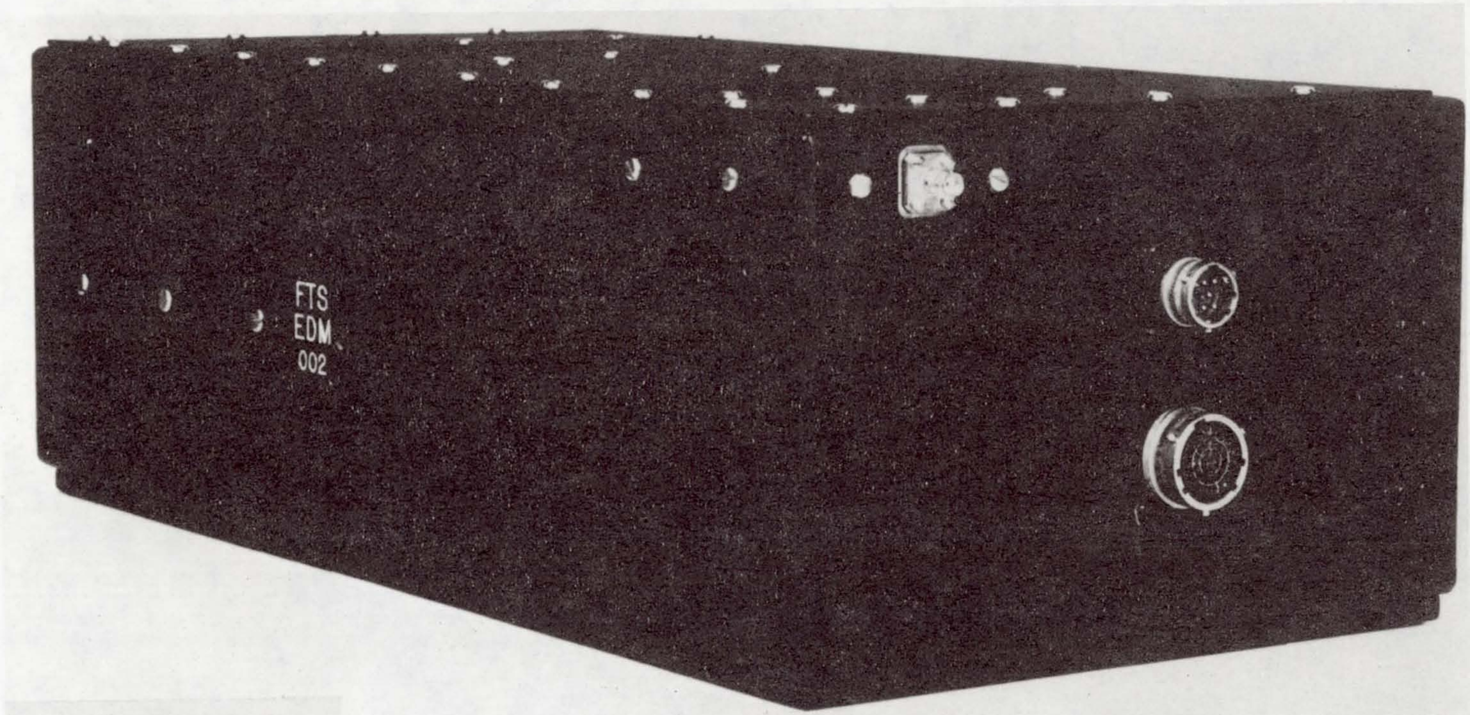


Figure 2

PHOTOGRAPH OF EDM

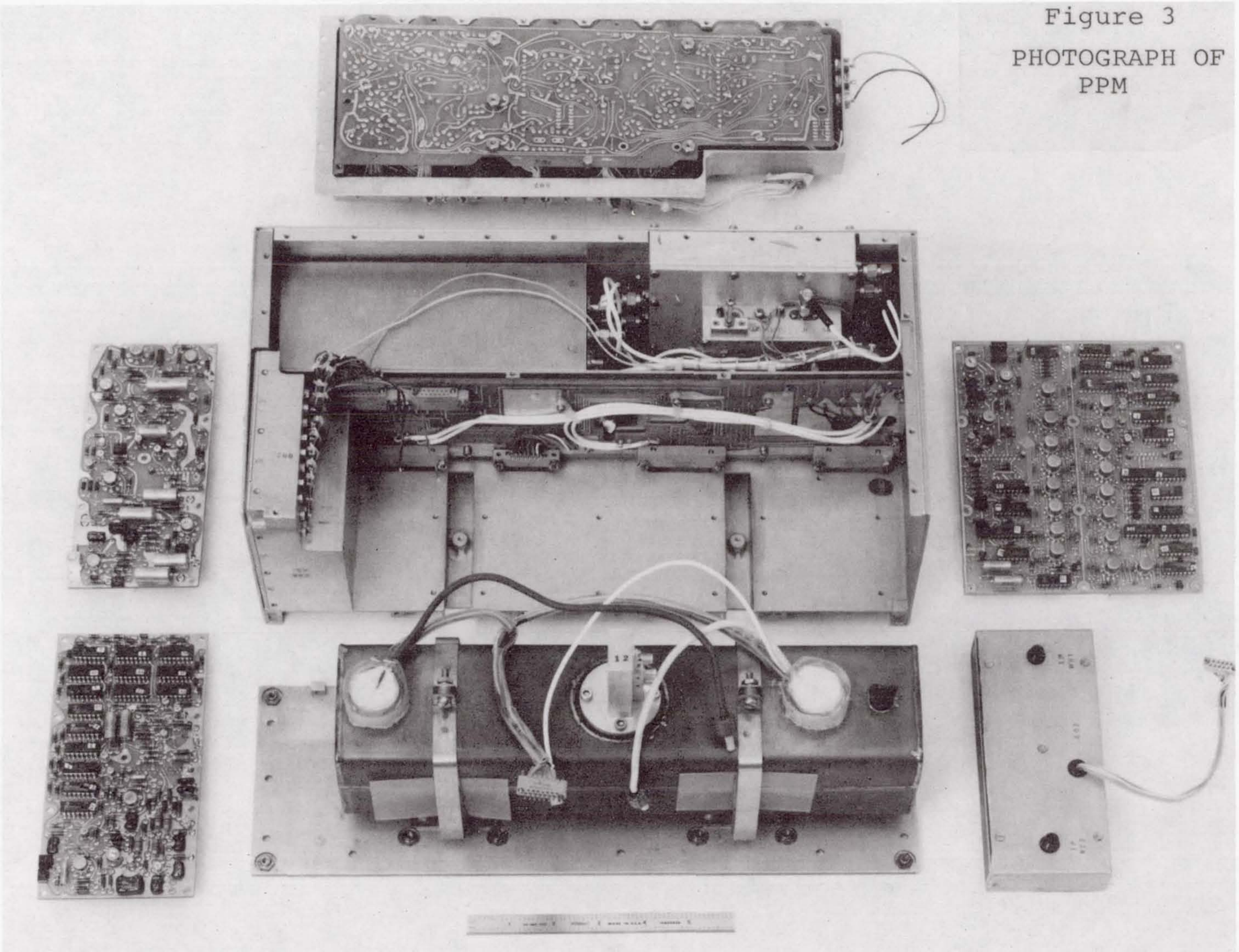


Figure 3
PHOTOGRAPH OF PPM

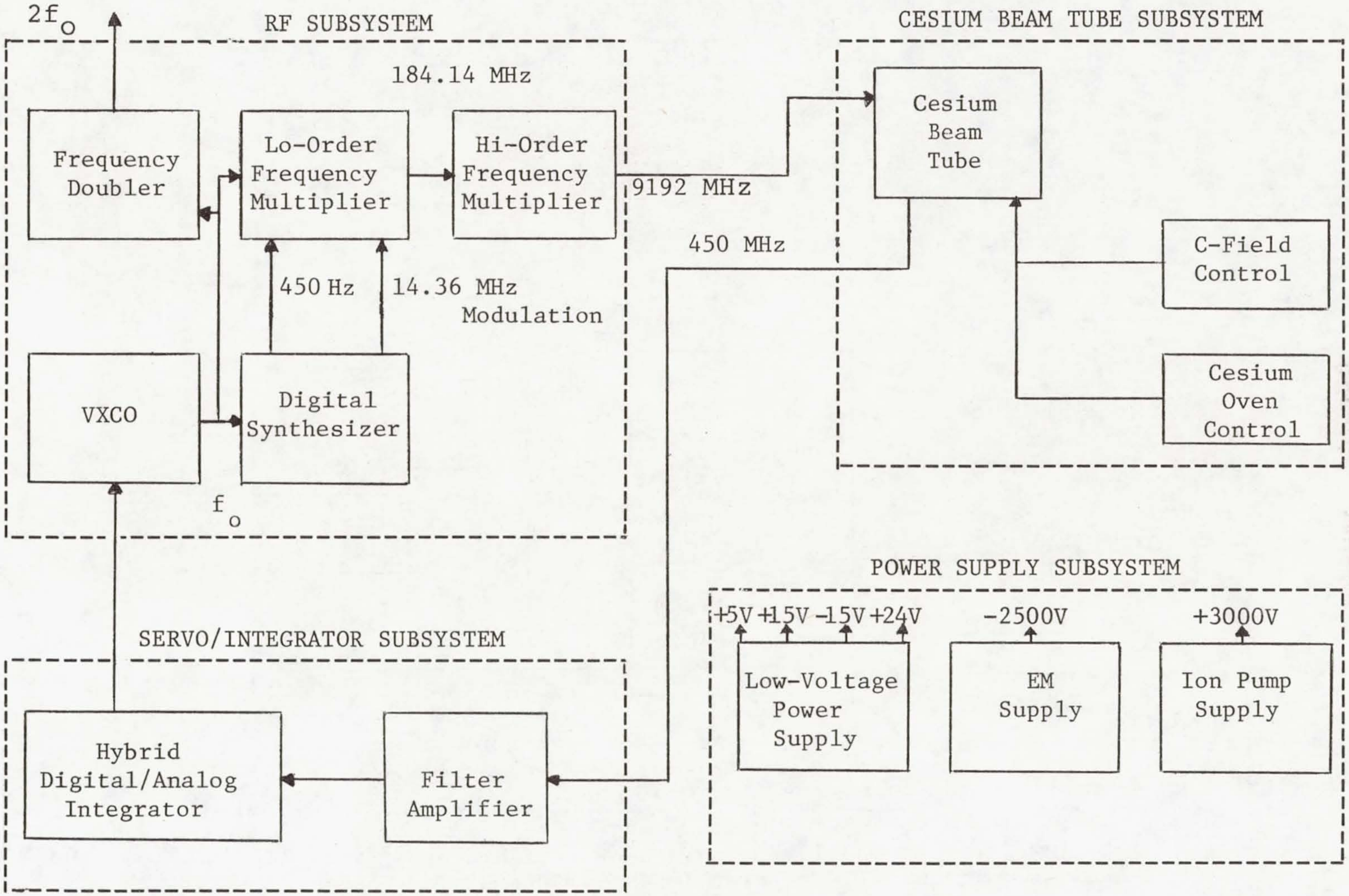


Figure 4

PPM CESIUM BEAM FREQUENCY STANDARD

SYSTEM BLOCK DIAGRAM

11/14/78 MWL

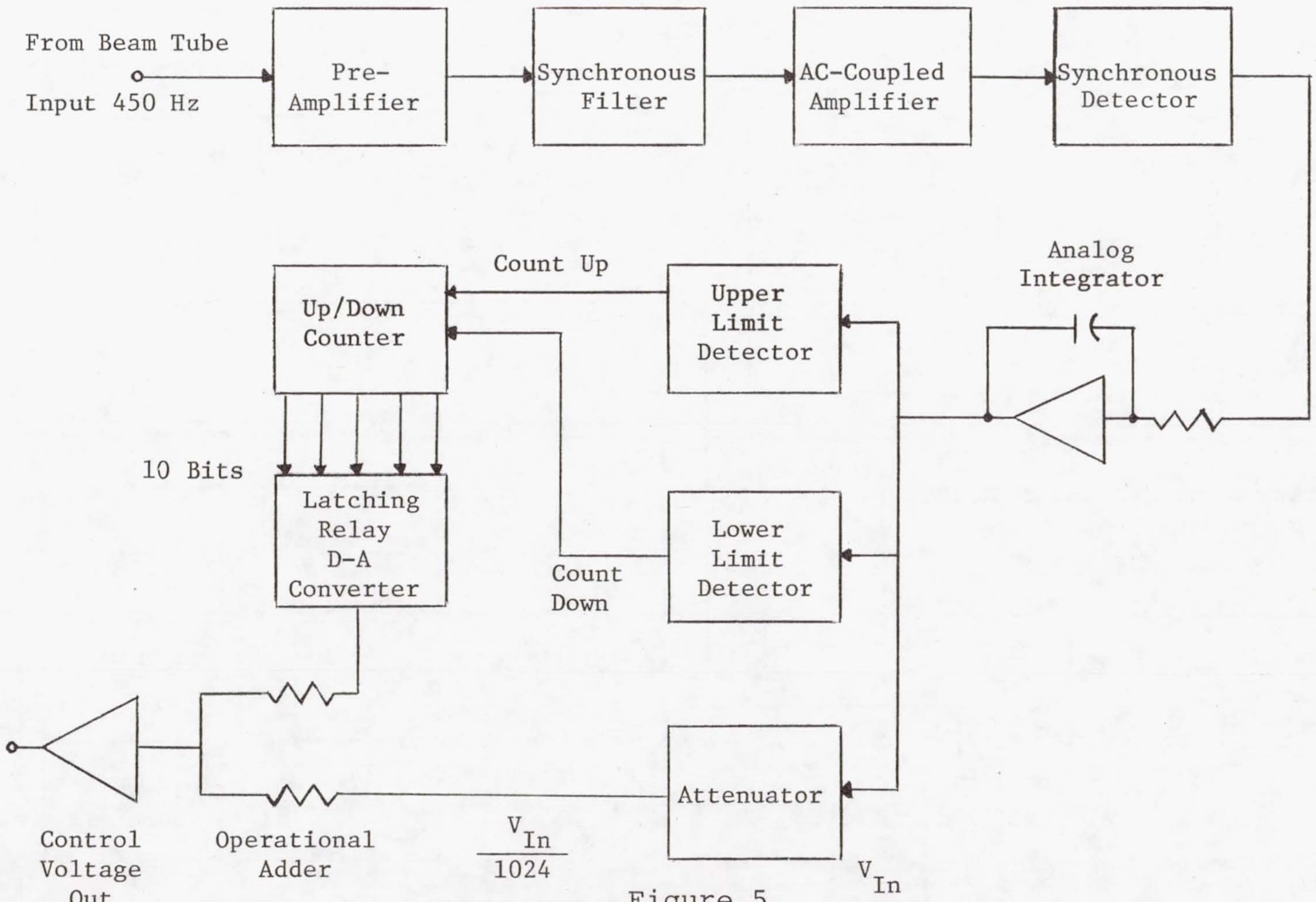


Figure 5
PPM SERVO/INTEGRATOR SUBSYSTEM
BLOCK DIAGRAM

11/14/78 MWL

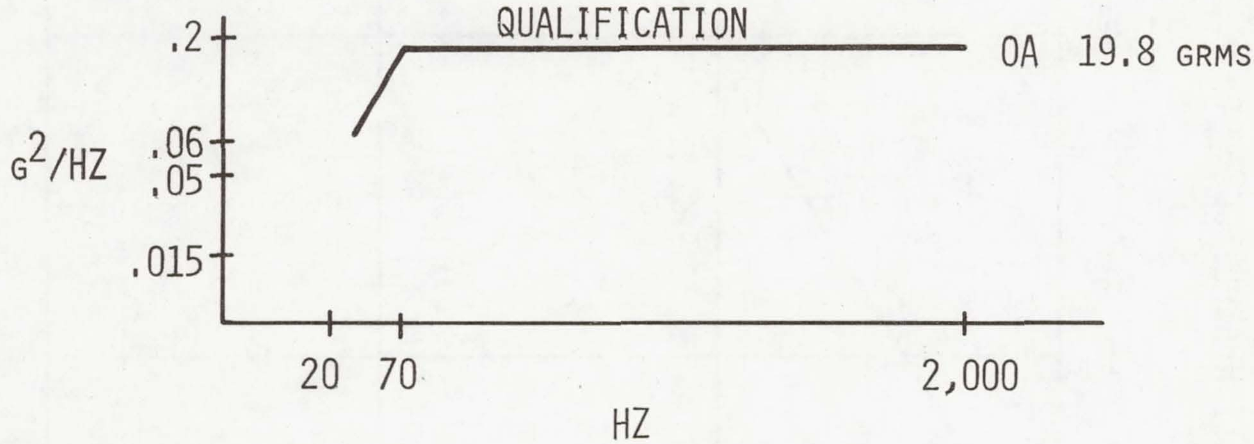


Figure 6. Vibration Level Requirements

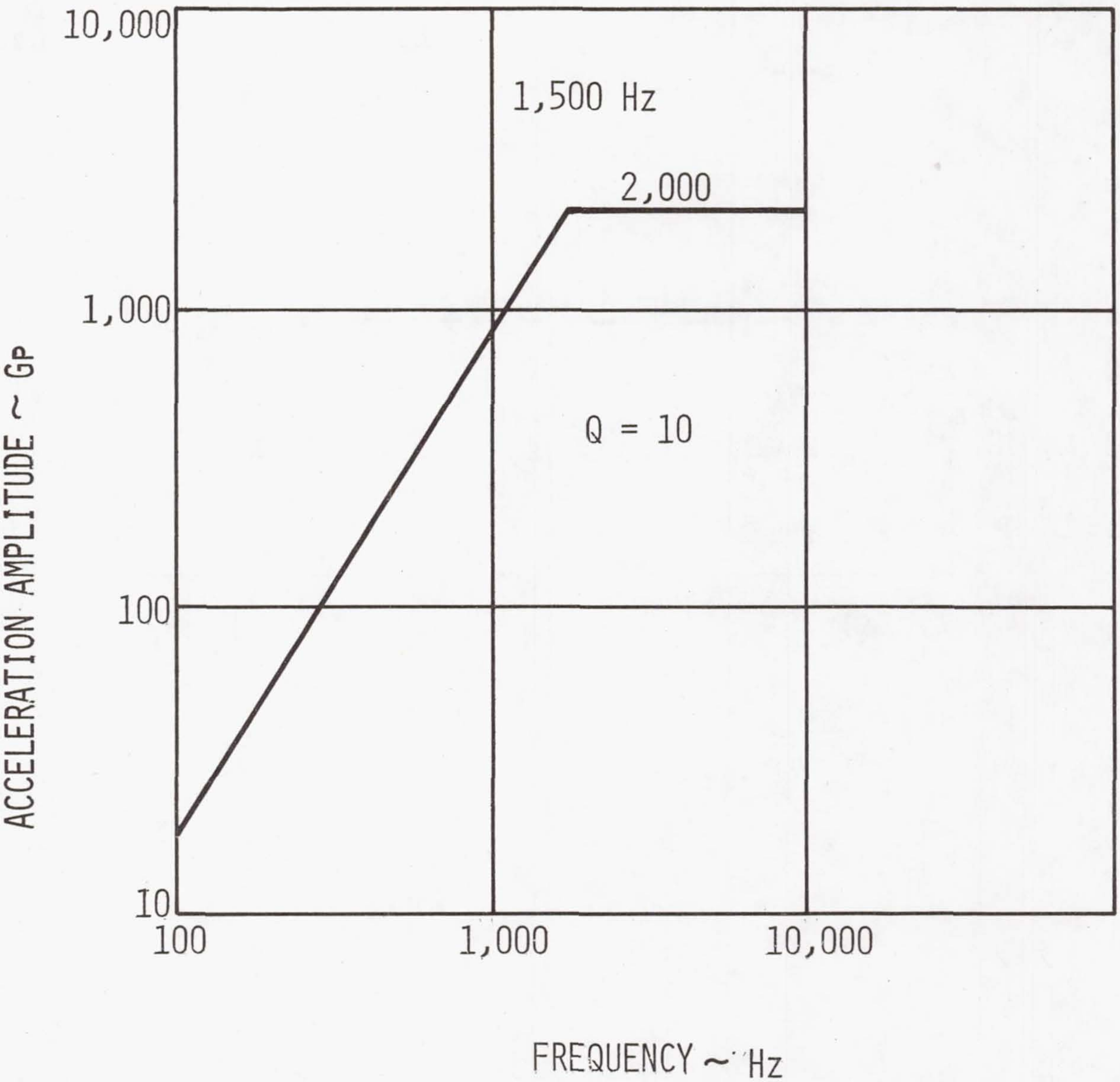


Figure 7. Shock Level Requirements for Qualification

T = 1 hour in air

T = 12 hours in vacuum

Figure 8

TEMPERATURE CYCLING REQUIREMENT FOR QUALIFICATION

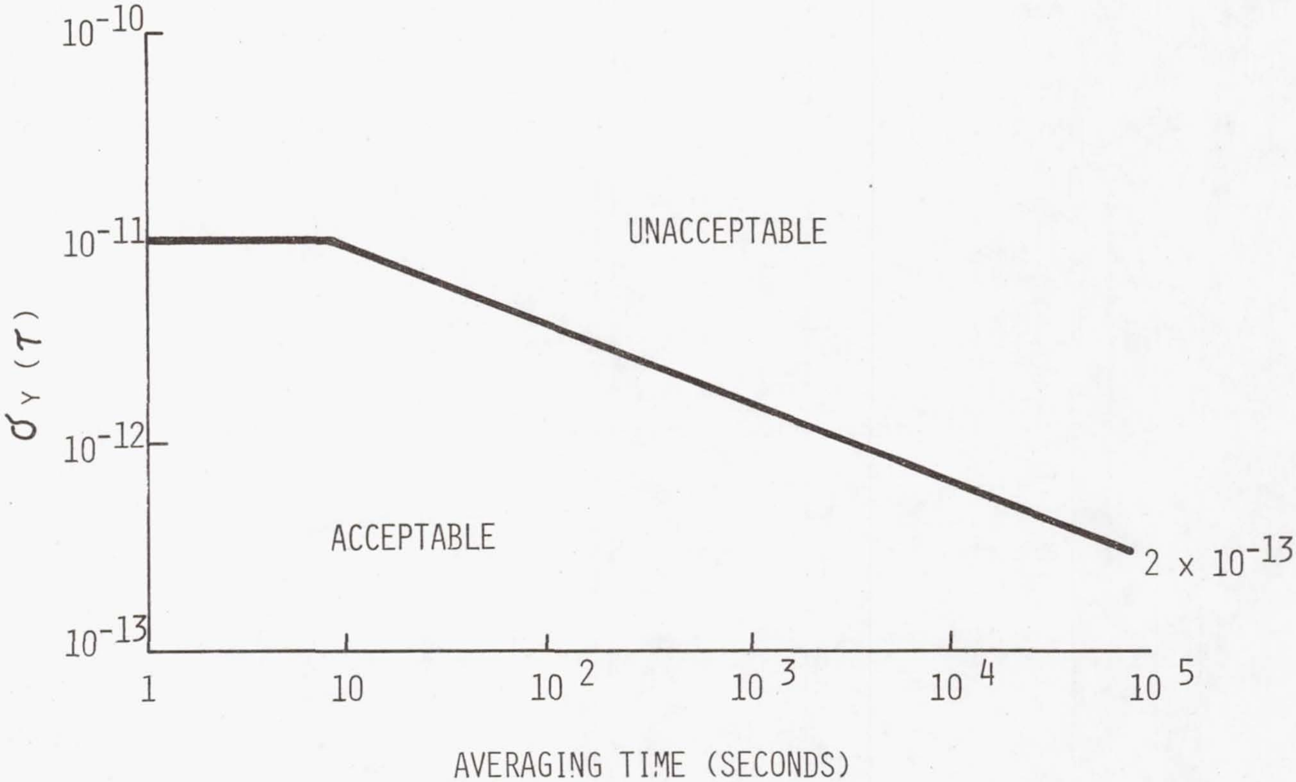


Figure 9. Frequency Stability Requirements

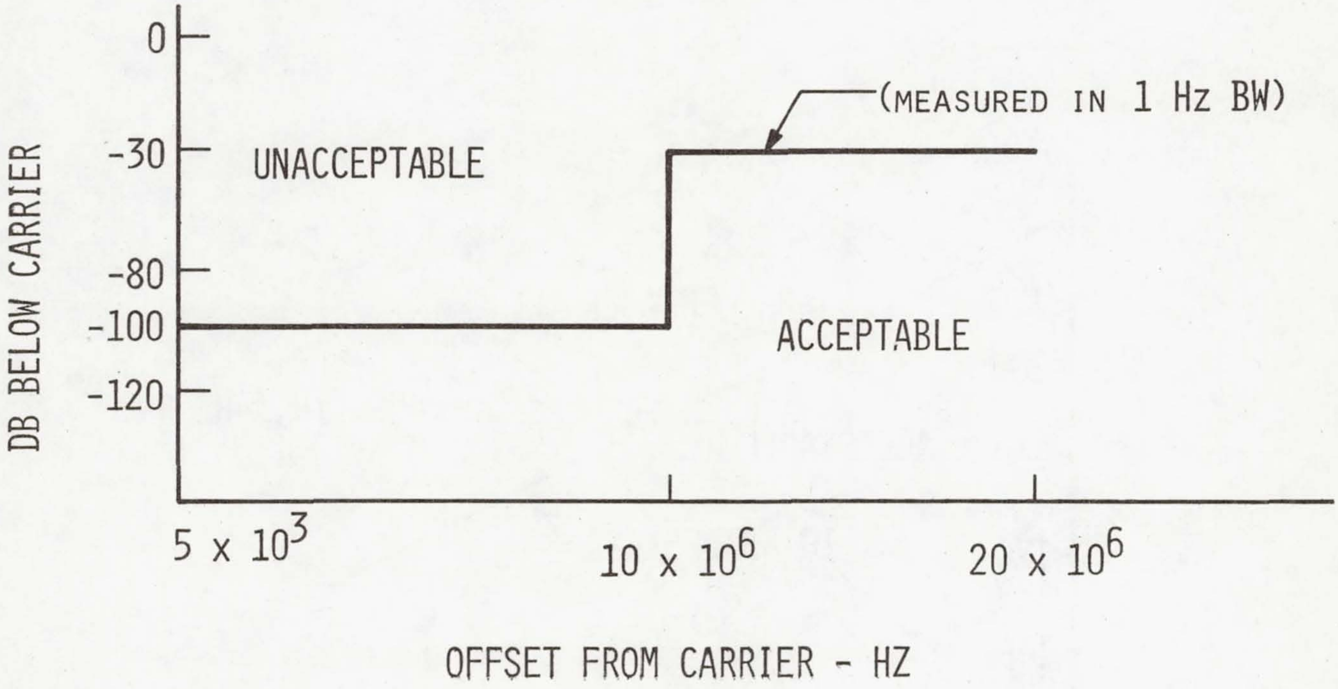


Figure 10. Spurious Signal Levels

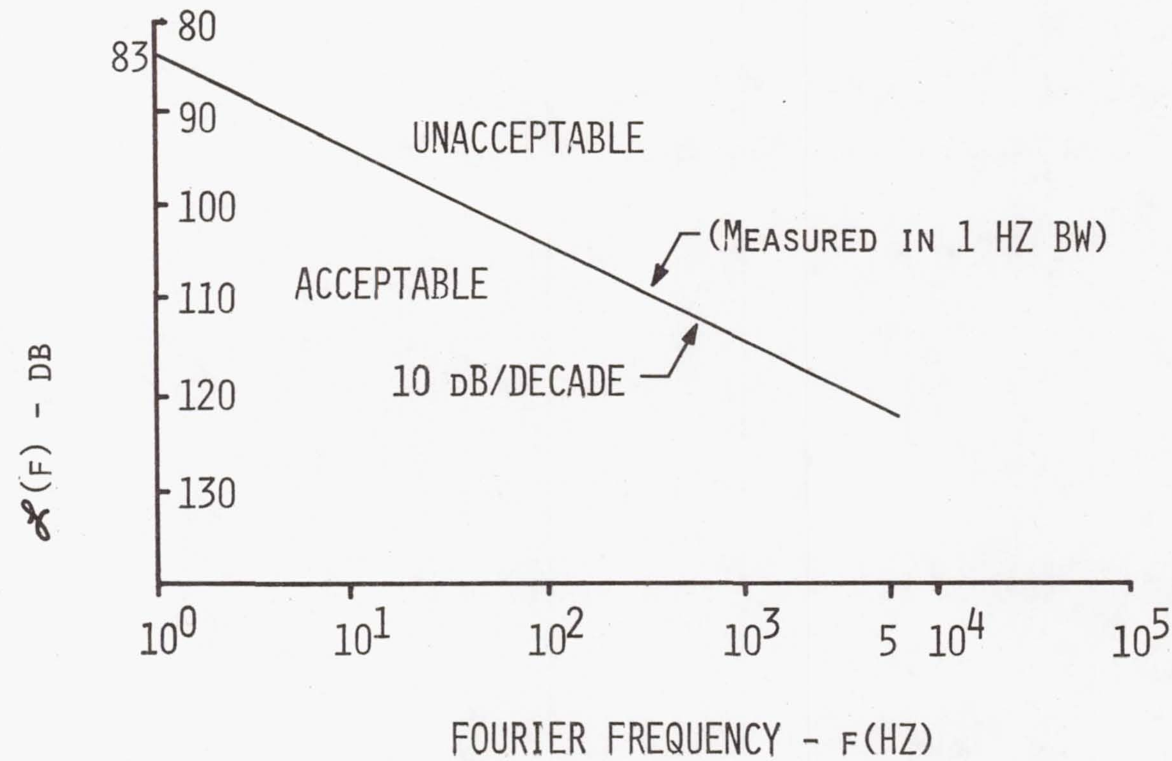


Figure 11
Phase Noise Versus Frequency Offset

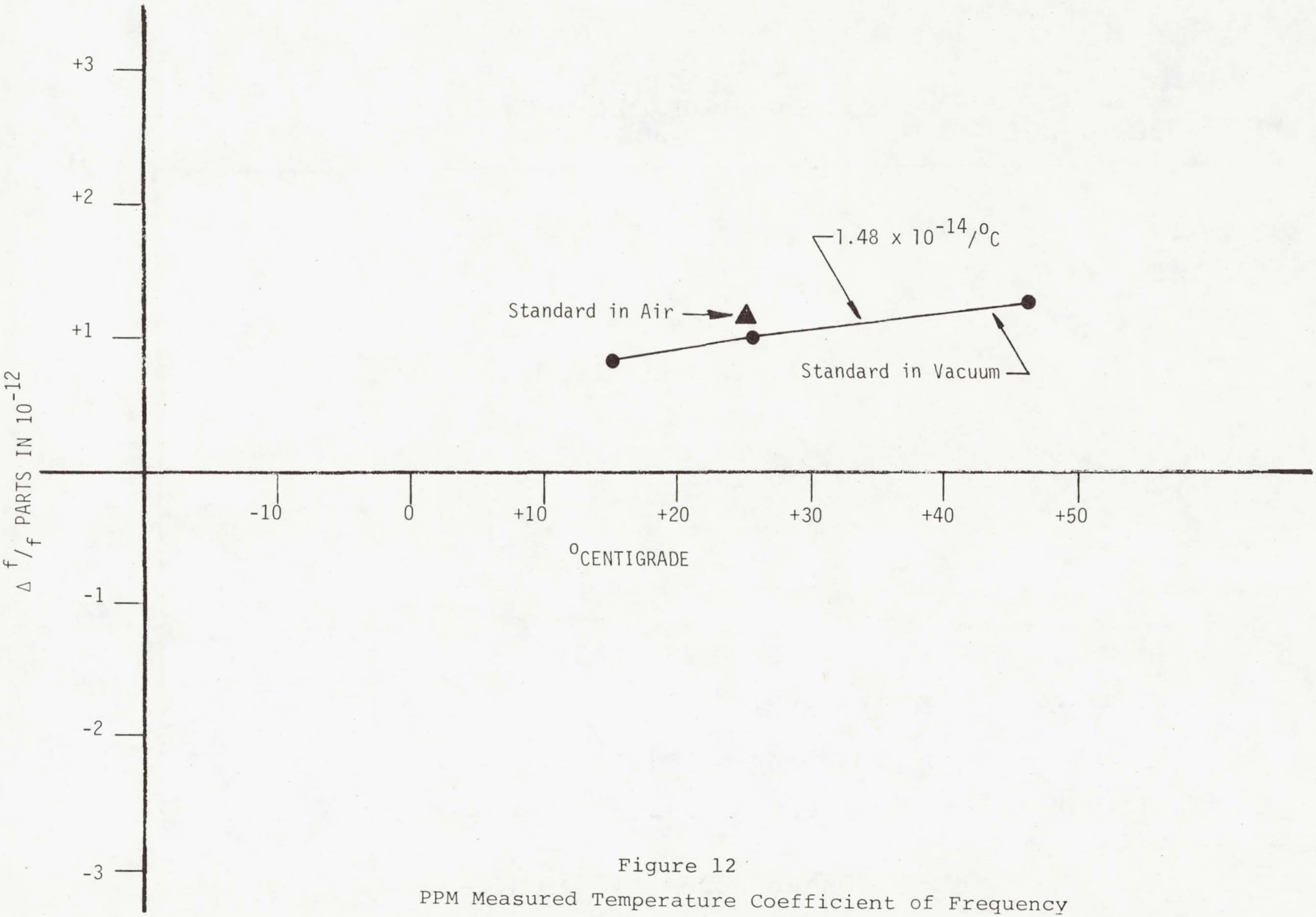


Figure 12
PPM Measured Temperature Coefficient of Frequency

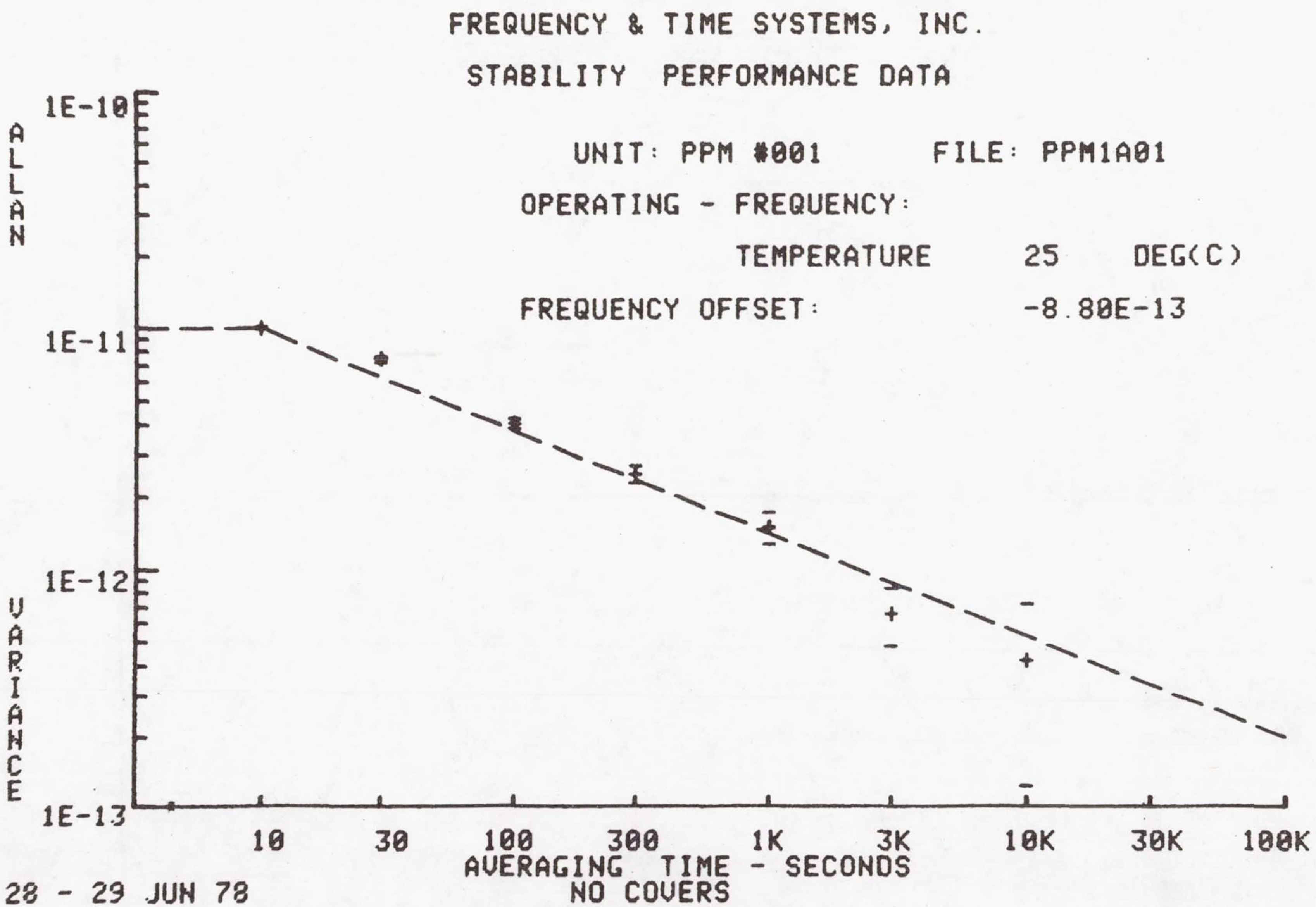


Figure 13

PPM Frequency Measured Stability

QUESTIONS AND ANSWERS

DR. HELMUT HELLWIG, National Bureau of Standards:

I have a comment. I am glad to see a portable clock finally, because I am getting tired of building my own--a portable clock in the sense of being really independent of power. You said eight hours?

DR. EMMONS:

Six hours. Two hours per pack. There are two battery packs, and, cleverly concealed under those is a last ditch pack spread out.

DR. GERNOT M. R. WINKLER, U.S. Naval Observatory:

Could you carry additional packs with you and plug them in, midterm, and prolong the clock's life if you see that you are stuck?

DR. EMMONS:

As each pack dies, you plug in another--before it dies.

MR. JECHART:

It would mean the purchase order for a lot of these--I would say we can do this order.

DR. EMMONS:

How many does he have?

DR. HELLWIG:

Efratom wants a purchase order also, if there are enough. A little more on the serious side; if there is a marketable product, I think somehow it will be built. And the technology for that (portable clocks) has existed for quite some time. Maybe we should ask FTS Company, "What do you expect as a market for this kind of device?" Don't answer if you don't want to.

DR. EMMONS:

I had better not. I don't feel up to that one.

DR. HELLWIG:

Maybe Dr. Winkler has a comment. No? No guess at the Market? You know the clock carrying business.

DR. WINKLER:

I think the answer depends entirely upon performance; I mean, performance under practical field conditions. And after we see that, we will say.

DR. EMMONS:

There has been a real snag, and that is in eliminating....well, since the advent of travel restrictions and anti-first class policies, we find a little bit of backlash concerning no more first class necessities.

DR. HELLWIG:

So that may create a market--no more first class travel. You are forced into economy or subeconomy, and then you have to buy little clocks that fit under the seat. Then they will be produced.

DR. WINKLER:

What is the price?

DR. EMMONS:

I would have to check from the floor here. Would someone care to comment on the price?

MR. THOMAS PARELLO, Frequency and Time Systems:

$26,000, and it is not GSA scheduled.

DR. EMMONS:

There is one back behind the screen if anyone wants to look at it.

MR. DAVID W. ALLAN, National Bureau of Standards:

A couple of questions, Don. On the stability slide, you showed just the white noise. Have you looked long enough to see if it starts to flatten and you get flicker noise or some other problems? And if so.....

DR. EMMONS:

My feeling is that we haven't looked long enough. I don't have any hot off the press data. I'm sorry.

MR. ALLAN:

The stability of 10^4 seconds is 4×10^{-13}, I believe you reported, and that seems inconsistent with the nominal 10 second data, as if it doesn't go as white noise. I was just curious whether you get some other strange thing there.

DR. EMMONS:

Well, I am inclined to say that maybe we have enough statistical spread to cover that.

MR. ALLAN:

To cover your tracks.

DR. EMMONS:

I will let you look in detail at it here.

DR. VICTOR REINHARDT, NASA Goddard Space Flight Center:

On the commercial units, since we have a little bit of interest in that, are the data that you got on the temperature coefficient, on the spacecraft unit, expected in the commercial unit? The 1.5 in 10^{14} per degree C?

DR. EMMONS:

I'm sorry. I don't have good numbers with me, but the commercial unit behaves very well over temperature. Do you remember some numbers, Tom?

MR. PARELLO:

Not precisely, but if I had to take a guess, I would say it is perhaps parts in 10^{12} over the full -28 to +61 degree C range.

DR. EMMONS:

Yes, over the full temperature range, parts in 10^{12}.

Page intentionally left blank

DISCUSSION FORUM: ATOMIC FREQUENCY STANDARDS

CESIUM

C. C. Costain, National Research Council of Canada

I am not sure how this panel was picked, but I expect it was for our objectivity. I will therefore try to be objective on the subject of cesium standards. They are certainly important to us in the standards laboratories because, I think, they will remain as the defined basis for time and frequency for at least the next twenty years.

I am not going to say much about commercial cesium standards. They are widely used and well known. They are normally within specifications, and most of us are annoyed if they do not perform an order of magnitude better than specifed. The entry of the Frequency and Time Systems into the field is important, and our measurements, and I think others, show the FTS performance is between the HP and HP high-performance option. My only question with respect to commercial standards would be as to whether too much lifetime is sacrificed in attaining short-term stability.

I am now going to go directly to a discussion of our primary cesium standards at NRC. I have a particular reason for doing so, which will become obvious. CsV has been operating continuously since May 1, 1975, and has undergone six full evaluations in that time. If it is assumed that TAI has been decreasing in frequency by 8×10^{-14}/year, with this one-parameter fit, the standard deviation between TAI and CsV, from the BIH circular D, is less than 0.4 μs. We do not know what the flicker floor of CsV is because we have nothing as good to measure it.

Al Mungall and Herman Daams have completed the three new standards, CsVI, A, B and C. The next two figures show part of the construction. Figure 1 shows the inner C field structure, and the six coils to measure the LF resonances. Figure 2 shows the three standards completed. They have been operating as clocks for a few weeks, but they have not been evaluated. This will take most of a year, but the resonances are beautifully symmetric out to the $m = -3$ and $m = +3$, with a symmetry which would delight any physicist.

But the stability to date has been disappointing. The Allan σ approaches 1×10^{-14}, and then after maybe 24 hours, a frequency change of up to 1×10^{-13} occurs. The culprit is the C field. Al Mungall has found by measuring the low frequency resonances that the change in frequency is the result of a change in the C field. Sometimes one, two or three of the coils show a change. It is the residual magnetism in the shields which is changing. Better degaussing is expected to reduce the effects, and work is proceeding on this feature.

Mungall has suggested that the magnetic shields could well be the limiting factor in the stability of atomic standards. The same effects occur in CsV, where changes of parts in 10^{15} are seen in the biweekly C field measurements. But the field of a dipole is dependent on the cube of the distance, and as CsVI is nearly a factor of 2 smaller diameter, the effects are nearly an order of magnitude larger. It is likely that the residual magnetism will have a much greater effect on the frequency than the distributed phase shift when the beam position is changed.

Perhaps in H-masers, when a dielectric cavity reduces the size by a factor of 3, magnetic effects which are now 3×10^{-15} could become one or two orders of magnitude larger. Certainly at this level of precision one must expect the unexpected, and it becomes increasingly difficult to convert dreams into reality.

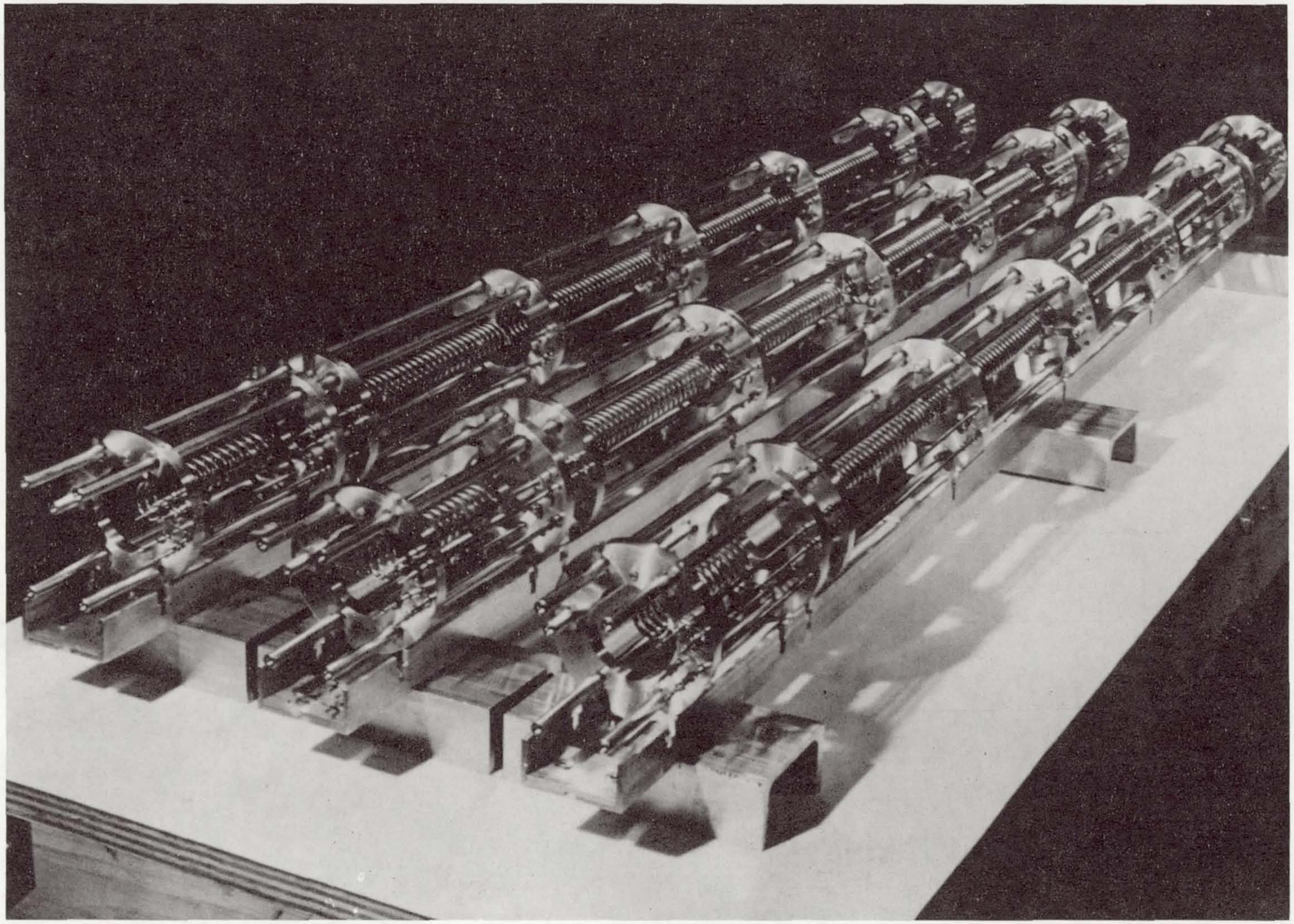


Figure 1. The C field structure of the three CSVI standards. The C field current in the four rods produces the transverse C field. The LF coils measure the C field outside each end of the cavity as well as in the drift space.

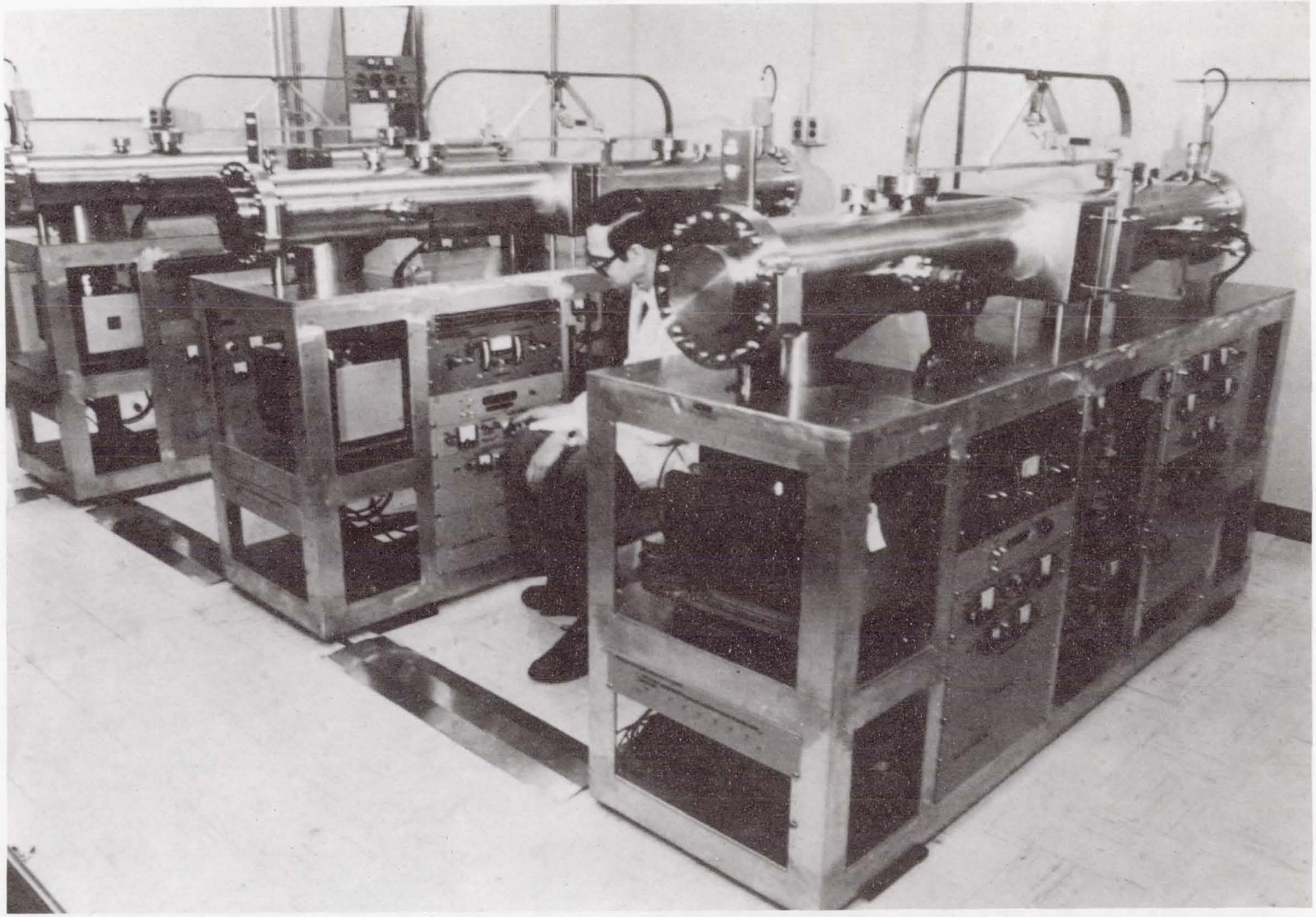


Figure 2. Dr. Mungall with the three CsVI primary clocks.

DISCUSSION FORUM: ATOMIC FREQUENCY STANDARDS

RUBIDIUM

Thomas C. English
Efratom Systems Corporation

INTRODUCTION

The purpose of this presentation is to point out the advantages of rubidium gas cell frequency standards relative to both quartz oscillators and other atomic standards. We also consider how these advantages determine the types of applications that are suitable for rubidium devices, and what improvements can be expected in the future.

MAJOR ADVANTAGES OF RB RELATIVE TO QUARTZ

We begin this presentation by enumerating the advantages of commercial rubidium frequency standards relative to commercial quartz oscillators. See Table 1. In the first column, the characteristic to be compared is listed. In the second column, values of these characteristics are given for a small commercial rubidium standard. All of the parameter values given in this column are realized <u>simultaneously</u> in a single, commercial device. In the third column, state-of-the-art parameter values are listed for <u>presently available commercial</u> quartz oscillators (developmental devices are not included!). It is important to point out here that these parameter values <u>cannot</u> be simultaneously realized in a single commercial quartz device, and that the values for a typical high quality commerical quartz oscillator are usually about an order of magnitude worse than shown here. For example, a typical, high quality commercial quartz oscillator will have drift rate of about 1×10^{-10}/<u>day</u>. The value of $< 2 \times 10^{-11}$/<u>day</u> indicated in Table 1 can be realized in a currently available quartz device, but the price tag is rather high, of the order of $15k. On the other hand, a long-term drift rate of less than 1×10^{-11}/<u>month</u> is readily available from a rubidium device. This is about a factor of 60 better than the table value of 2×10^{-11}/<u>day</u> for the best commercial quartz.

In summary, Table 1 shows that rubidium is one to two orders of magnitude better in each parameter listed, with the possible exception of short-term stability over periods of minutes to hours. Moreover, all parameter values given here are simultaneously realized in a small

TABLE 1

MAJOR ADVANTAGES OF RUBIDIUM RELATIVE TO

QUARTZ OSCILLATORS

CHARACTERISTIC	SMALL COMMERCIAL RUBIDIUM	STATE-OF-THE-ART PARAMETERS COMMERCIAL QUARTZ[A]
SHORT-TERM STABILITY (MINUTES TO HOURS)	PARTS IN 10^{13}	NOT USUALLY SPECIFIED
LONG-TERM DRIFT	$< 1 \times 10^{-11}$/<u>MONTH</u>	$< 2 \times 10^{-11}$/<u>DAY</u>
WARMUP TIME (25 °C AMBIENT)	10 MIN TO $< 2 \times 10^{-10}$	30 MIN TO 1×10^{-9}
RETRACE (ON-OFF 24 HRS-ON)	$< 2 \times 10^{-11}$	1×10^{-9}
ACCELERATION SENSITIVITY	$< 8 \times 10^{-12}$/G	8×10^{-10}/G

[A] THESE PARAMETERS ARE NOT SIMULTANEOUSLY AVAILABLE IN A SINGLE DEVICE.

commercial rubidium, whereas this is not the case for commercial quartz.

MAJOR ADVANTAGES OF RB RELATIVE TO OTHER ATOMIC STANDARDS

The main advantages of rubidium relative to other atomic standards are listed in Table 2. These advantages include small size, light weight, low power consumption and low cost.

PHOTOGRAPH OF SMALL COMMERCIAL RB

Figure 1 shows the size of a small commercial rubidium frequency standard. It is a cube that is 4 inches on a side. The pocket watch serves to give one a gut feel for the small size of this device.

PHYSICAL CHARACTERISTICS OF COMMERCIAL RB & CS STDS

The first two lines of Table 3 compare small commercial rubidium and cesium devices. These are the basic, no-frills units. Note that rubidium is 8 times smaller, 7 times lighter, uses ½ as much power and costs from 1/3 to 1/5 as much.

The last two lines of Table 3 are for those persons who are interested in a bench or rack mount unit, including an AC power supply and a standby battery pack for uninterrupted operation in the event of a powerline failure. In this case, small size, weight and power consumption are not of major concern, so no effort has been made to minimize these characteristics.

By the way, hydrogen devices have not been included in this comparison because we are concerned here only with commercially available atomic standards; to the best of our knowledge, there are no commercially available hydrogen devices.

SIZE COMPARISON OF TWO COMMERCIAL ATOMIC STANDARDS

Figure 2 allows a direct comparison of the relative sizes of a small commercial rubidium and a small commercial cesium. For many years I did physics research in the area of atomic and molecular beams, with big, long, machines that filled up most of a room. For this reason, it is always amazing to me to see that it has been possible to make cesium standards as small as they are today. But, of course, the same is also true for present-day rubidium devices. In any case, it

TABLE 2

MAJOR ADVANTAGES OF RUBIDIUM RELATIVE TO

OTHER ATOMIC STANDARDS

- SMALL SIZE
- LIGHTWEIGHT
- LOW POWER CONSUMPTION
- GOOD SHORT-TERM STABILITY
- LOW PHASE NOISE
- POTENTIALLY FASTER WARMUP
- LOW COST

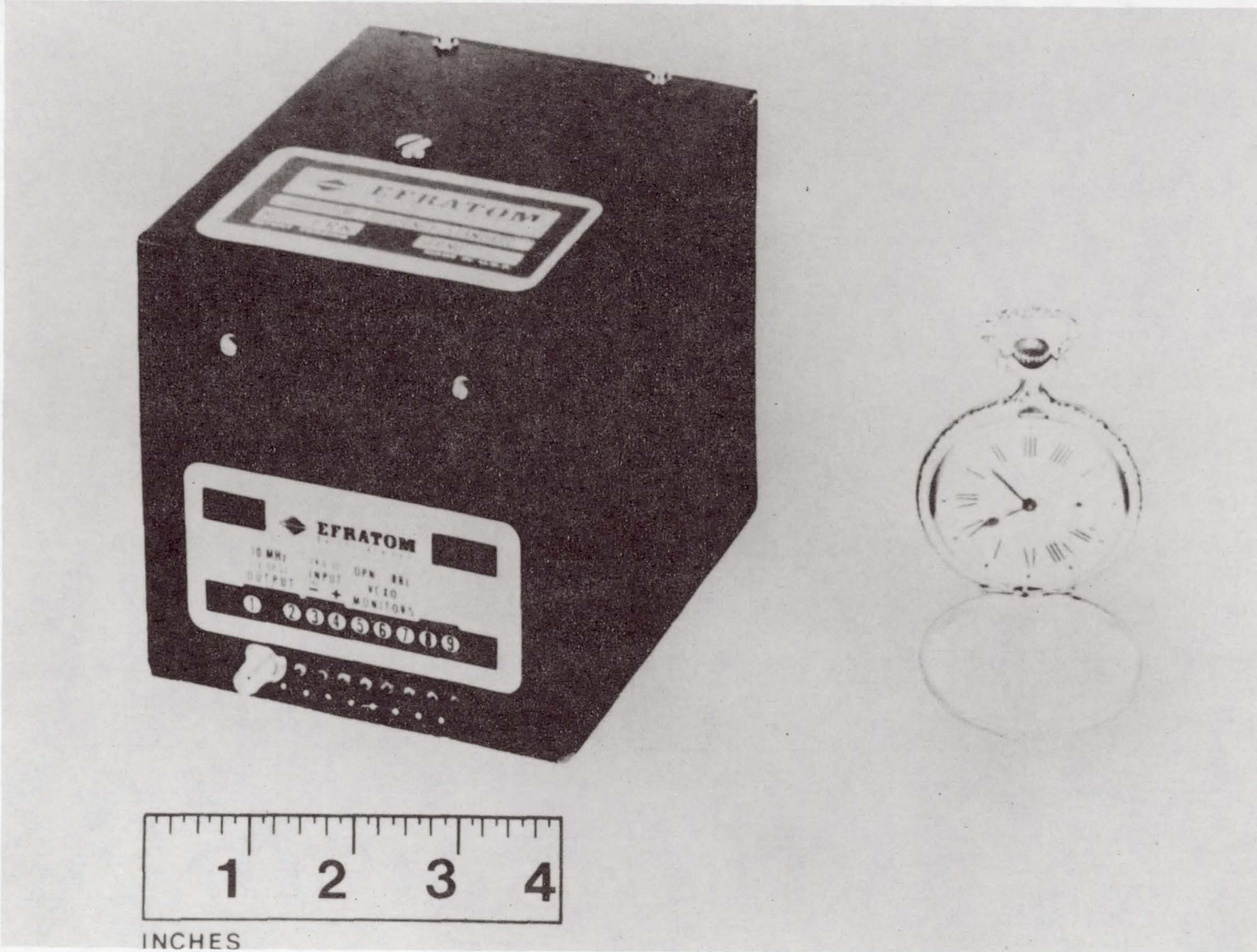


Figure 1. Small commercial rubidium frequency standard.

TABLE 3

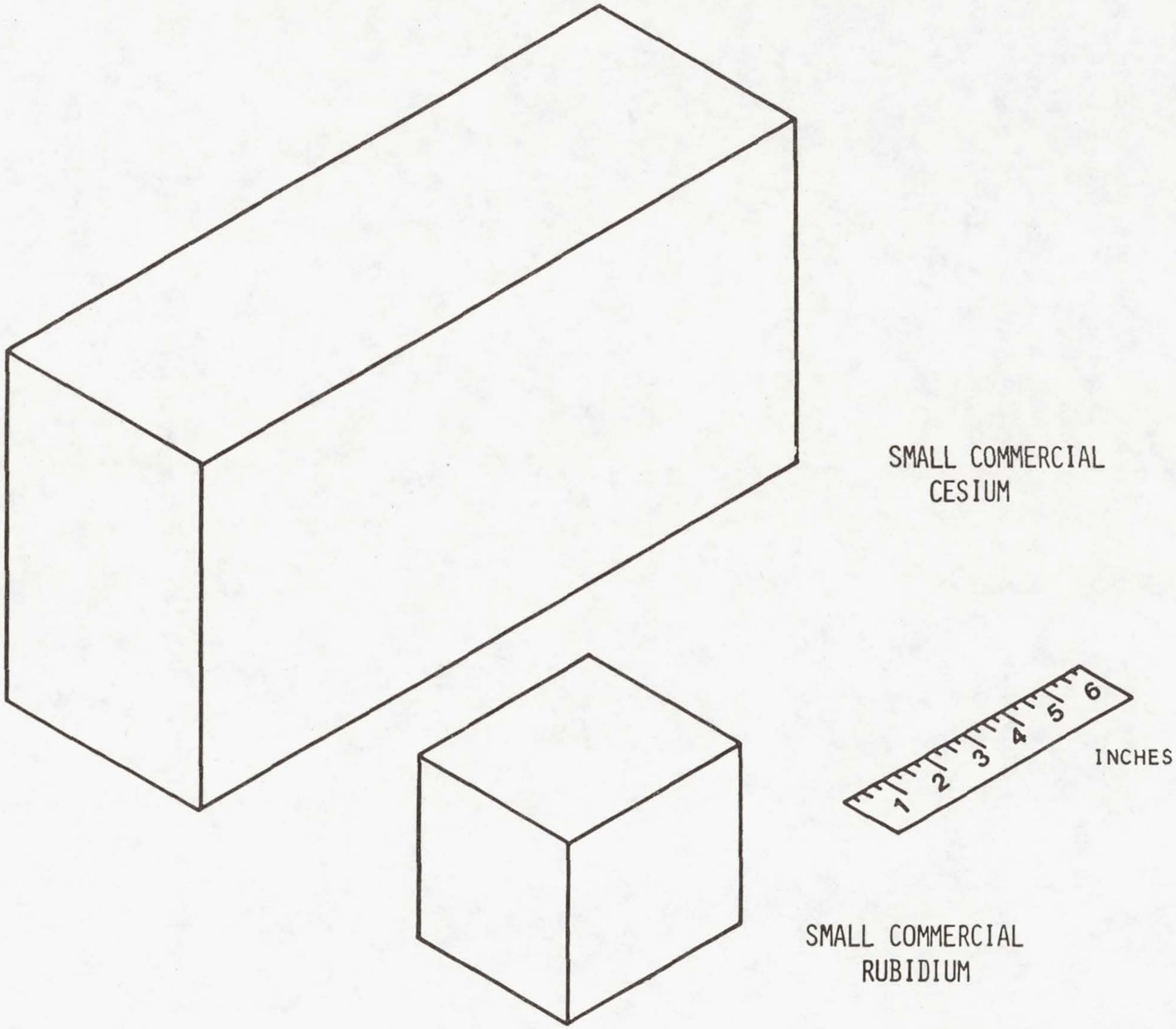
PHYSICAL CHARACTERISTICS OF COMMERCIAL RUBIDIUM AND
CESIUM FREQUENCY STANDARDS

DEVICE	SIZE (CU IN)	WEIGHT (LBS)	DC POWER (W)	COST[B] (K$)
SMALL COMMERCIAL RUBIDIUM	67	3	13	4 - 6
SMALL COMMERCIAL CESIUM	560	22	24	19
RUBIDIUM/CESIUM	x 1/8	x 1/7	x 1/2	x 1/3 - 1/5
COMMERCIAL RUBIDIUM [A]	900 - 1400	27 - 38	---	7 - 10
COMMERCIAL CESIUM [A]	1600 - 2400	57 - 75	---	20 - 25

[A] INCLUDES AC POWER SUPPLY & STANDBY BATTERY PACK.

[B] SINGLE QUANTITIES.

FIGURE 2. SIZE COMPARISON OF TWO COMMERCIAL ATOMIC STANDARDS



is evident from Figure 2 that the small size of rubidium devices is one of their major advantages.

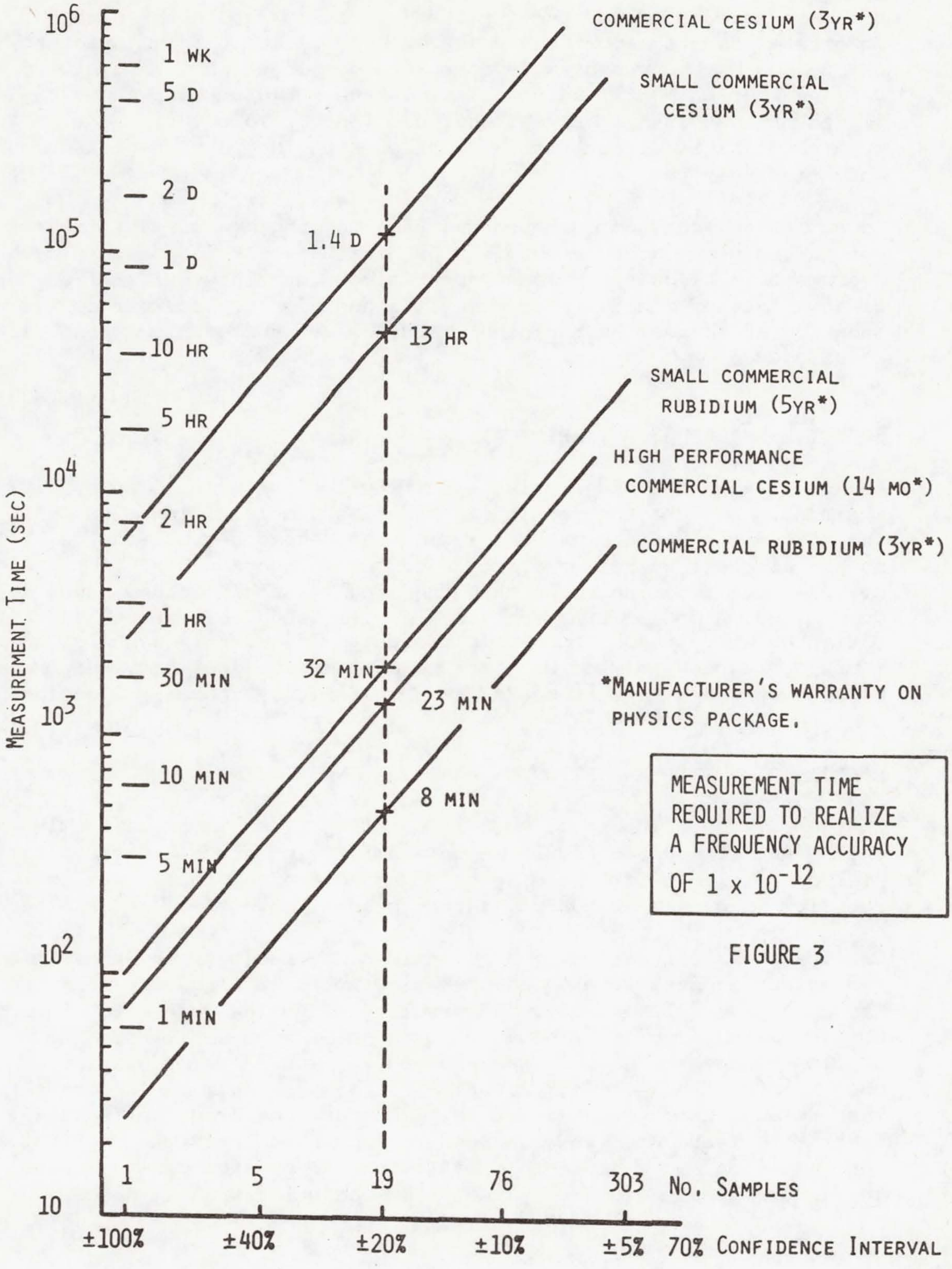
MEASUREMENT TIME REQ'D TO REALIZE A FREQUENCY ACCURACY OF 1×10^{-12}

Figure 3 shows the time required to make a frequency measurement to 1 part in 10^{12} using commercially available atomic frequency standards. In plotting these curves, we have assumed that the performance of the frequency standard used as the measuring device is the limiting factor. For each curve, the measuring device is specified to the right of the curve; for example, the measuring device for the uppermost curve is a commercial cesium.

In general, it should be obvious that the measurement time will depend on how good the short-term stability of the measuring device is; the better the short-term stability, the shorter the measurement time that is required. Because of the excellent short-term stability of rubidium standards, the measurement time required to attain 1×10^{-12} accuracy with them is very short.

In discussing measurement time, it is important to understand that we are really dealing with frequency fluctuations over a given period of time, and that these fluctuations are statistical in nature. For this reason, it is necessary to make multiple measurements in order to reduce the statistical uncertainty. For example, 19 measurements are required to specify an accuracy of 1×10^{-12} to within ± 20 %. For the commercial rubidium having the best available short-term stablity, this will require a total measurement time of 8 minutes. For the small commercial rubidium, 32 minutes will be required. When we look at the measurement times for the cesiums, we can see how good the rubidium times really are. One commercial cesium, a very commonly used one, requires a total measurement time of 1 day! A small commercial cesium is available that requires only about half this amount of time, but this is still quite long when compared to the rubidium figures.

Now, you will look at this graph and say, "but hey, wait, you forgot one of the cesiums!--the high performance cesium." Yes, you are right, the high performance cesium has good short-term stability -- it is comparable to that of the rubidiums, but it is obtained at a price. It is obtained by increasing the cesium beam intensity by more than an order of magnitude, and this reduces the life of the beam tube. This reduced lifetime is reflected in the manufacturer's warranty for the beam tube. For most commercial cesiums, the warranty is 3 years, but for the high performance cesium, the warranty is only 14 months. Here again, the cost factor enters: in general,



MEASUREMENT TIME REQUIRED TO REALIZE A FREQUENCY ACCURACY OF 1×10^{-12}

FIGURE 3

the beam tube replacement costs for any cesium are on the order of, or greater than the purchase price of a complete rubidium frequency standard. In the rubidium devices, the component in the physics package that is most likely to fail is the rubidium lamp whose replacement cost is only a few hundred dollars. Moreover, the manufacturers' warranties vary from 3 to 5 years on the physics package, which includes the lamp.

To summarize, conventional cesiums require long measurement times to attain frequency accuracies of 1 part in 10^{12}. It is possible to buy cesiums that allow short measurement times, but they suffer from the disadvantage of reduced beam tube life and high replacement costs. Rubidium standards, on the other hand, do not suffer from these disadvantages.

RUBIDIUM PHASE NOISE SPECIFICATION

Figure 4 shows the phase noise specification for a small commercial rubidium. Low phase noise is important when multiplying signals in the MHz region up into the GHz region and beyond because the noise power increases by n^2 for a frequency multiplication by a factor of n. The specification shows that the single sideband phase noise is down by 92 dB one Hz away from the carrier, and decreases as $1/f^3$ until the white phase modulation floor of -155 dB is reached at a Fourier frequency of 100 Hz. To the best of my knowledge, the phase noise spec shown here is better than that of any commercial cesium.

EFFECT OF NUCLEAR RADIATION ON AN OPERATING RB STD

One topic, about which not much information seems to be available, is the effect of nuclear radiation on atomic frequency standards. Data are now available for the effects of dose rate and total dose on rubidium frequency standards, and we present some of these data here.

Table 4 shows the result of a recently conducted test to determine the effect of dose rate on an operating rubidium frequency standard. The unit tested is one of the Rockwell engineering models for the GPS satellite program. This unit uses an Efratom small rubidium physics package. The unit was exposed to flash x-ray radiation at a dose rate of about 4×10^8 rads/sec while operating. This dose rate was the maximum dose rate that could be obtained from the flash x-ray facility. There are two main results from this experiment. First, the radiation had a negligible effect on the physics package. Second, the accumulated phase error due to the radiation was < 1 nsec.

FIGURE 4
PHASE NOISE
SMALL COMMERCIAL RUBIDIUM

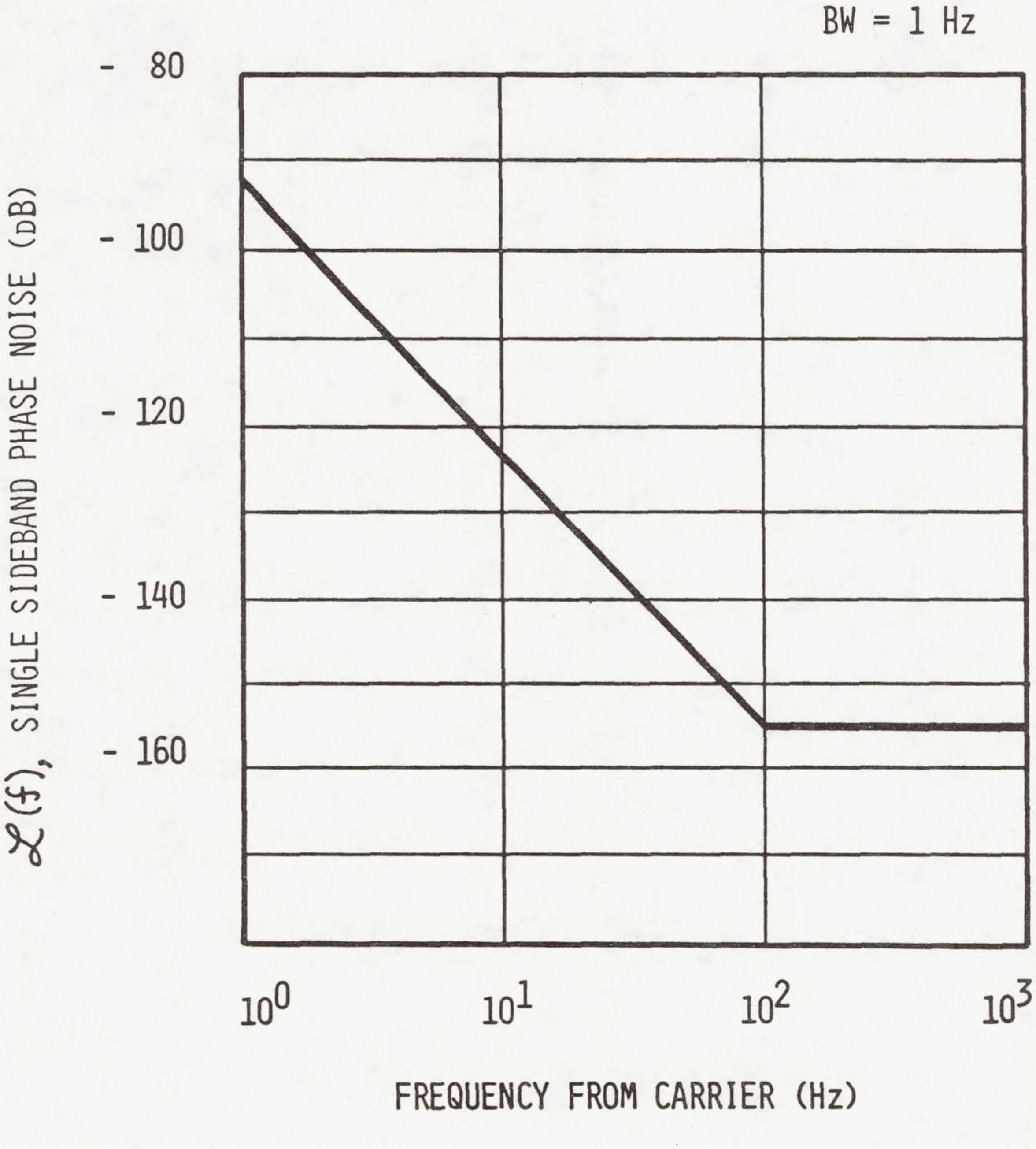


TABLE 4

EFFECT OF NUCLEAR RADIATION (DOSE RATE) ON AN OPERATING RUBIDIUM FREQUENCY STANDARD

DEVICE TESTED : UNSHIELDED GPS PHASE I RUBIDIUM SPACE CLOCK (EM3)

LOCATION OF TEST : ROCKWELL INTERNATIONAL, AUTONETICS DIVISION, FLASH X-RAY FACILITY

DOSE RATE : $> 3.8 \times 10^8$ RAD (Si)/SEC (MAX ATTAINABLE RATE)

RESULTS

- EFFECT OF RADIATION ON PHYSICS PACKAGE: NEGLIGIBLE
- ACCUMULATED PHASE ERROR: < 1 NSEC

ACCUMULATED PHASE ERROR FOR MOST SENSITIVE DIRECTION

The engineering model tested contained two radiation-hardened crystal oscillators (VCXO's). The first VCXO was used in the primary loop and was locked to the rubidium resonance with a loop time constant of < 0.1 sec. The second VCXO was used in a secondary 10.23 MHz loop that was locked to the first loop with a time constant of 21 sec. The main effect of the radiation is to alter the properties of the radiation-hardened VCXO's. This results in VCXO frequency changes which are subsequently servoed out by the control loops (each VCXO is locked, in effect, to the rubidium resonance). However, accumulated phase changes will result if the VCXO frequency changes occur in times short compared to the loop time constant; i.e., transient effects are responsible for the accumulated phase errors.

Figure 5 shows the accumulated phase error for the secondary loop. The radiation burst occurred at $t = 0$ while the unit was operating. After about 1 minute the phase stabilized with an accumulated phase error of about 22 nsec. Under the same conditions, the accumulated phase error for the primary loop was < 1 nsec. This difference may be attributed mostly to the smaller time constant for the primary loop and the fact that the rubidium resonance is essentially unaffected by the radiation. Here the important results are those for the primary loop. Secondary loops are rarely used, and in any event can be considered as a loop that is external to the actual rubidium device, whereas the primary loop is part of the rubidium device.

EFFECT OF 10^4 RADS ON AN OPERATING, UNMODIFIED SMALL COMMERCIAL RUBIDIUM STANDARD.

Figure 6 shows the effect of total radiation dose from a cobalt 60 source on an operating, unmodified, small commercial rubidium standard, essentially an Efratom Model FRK with high reliability electronic components. That is, the device was unmodified in any essential respect as far as its capacity to resist radiation was concerned. The total dose of 10^4 rads was accumulated at a steady rate over a 1 hour period.

As a result of the irradiation, the frequency of the unit increased by about 6 parts in 10^{11}. This frequency change resulted from a change in the characteristics of the electronics in the servo loop. The photocell voltage, here labelled "Rb lamp voltage," changed by less than 1 %. This shows that the rubidium lamp and the physics package optics were essentially unaffected by the radiation. On the other hand, the VCXO control voltage changed by 6 volts, indicating that the VCXO characteristics had been altered by the radiation.

FIGURE 5
ACCUMULATED PHASE ERROR FOR MOST
SENSITIVE DIRECTION

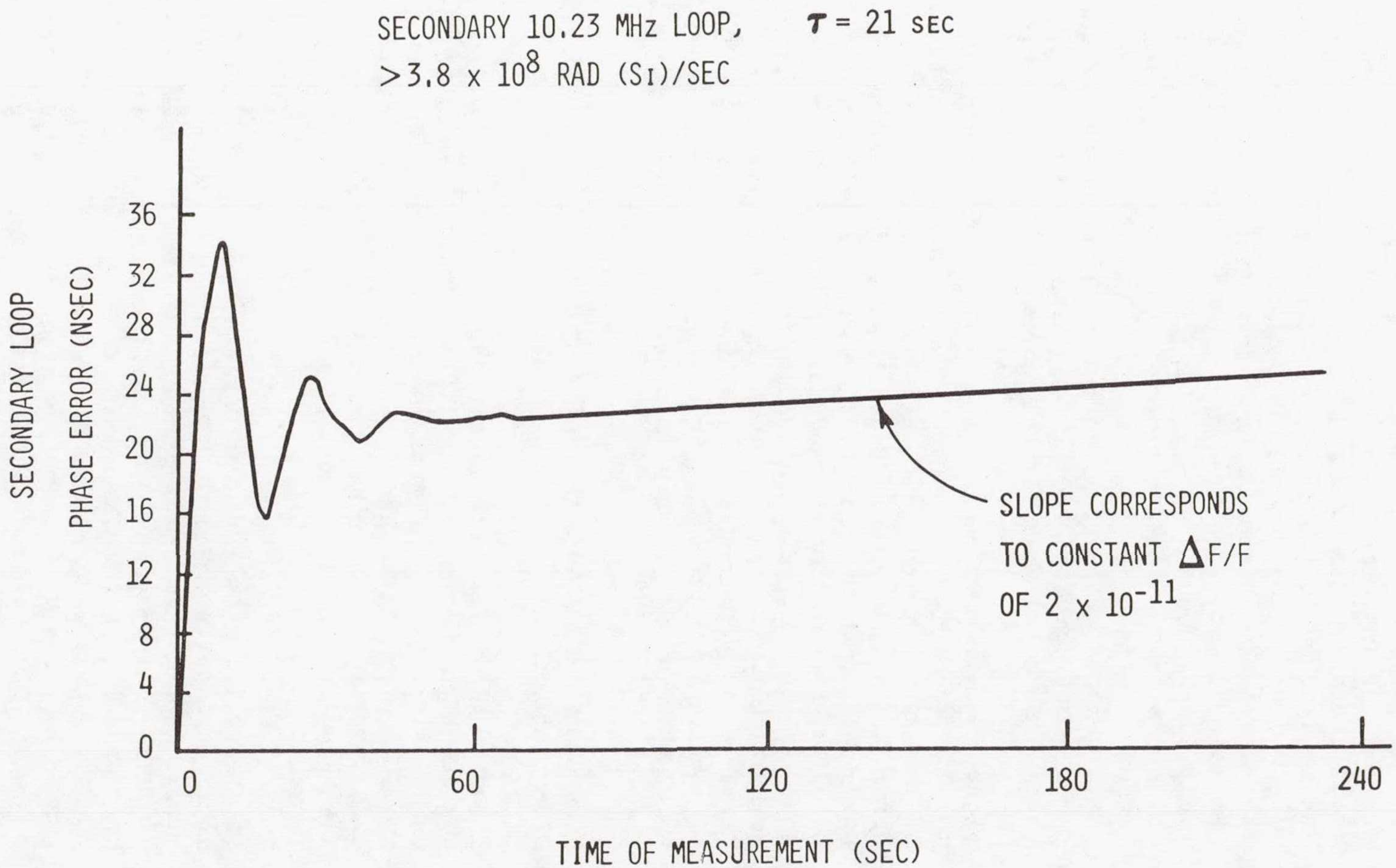
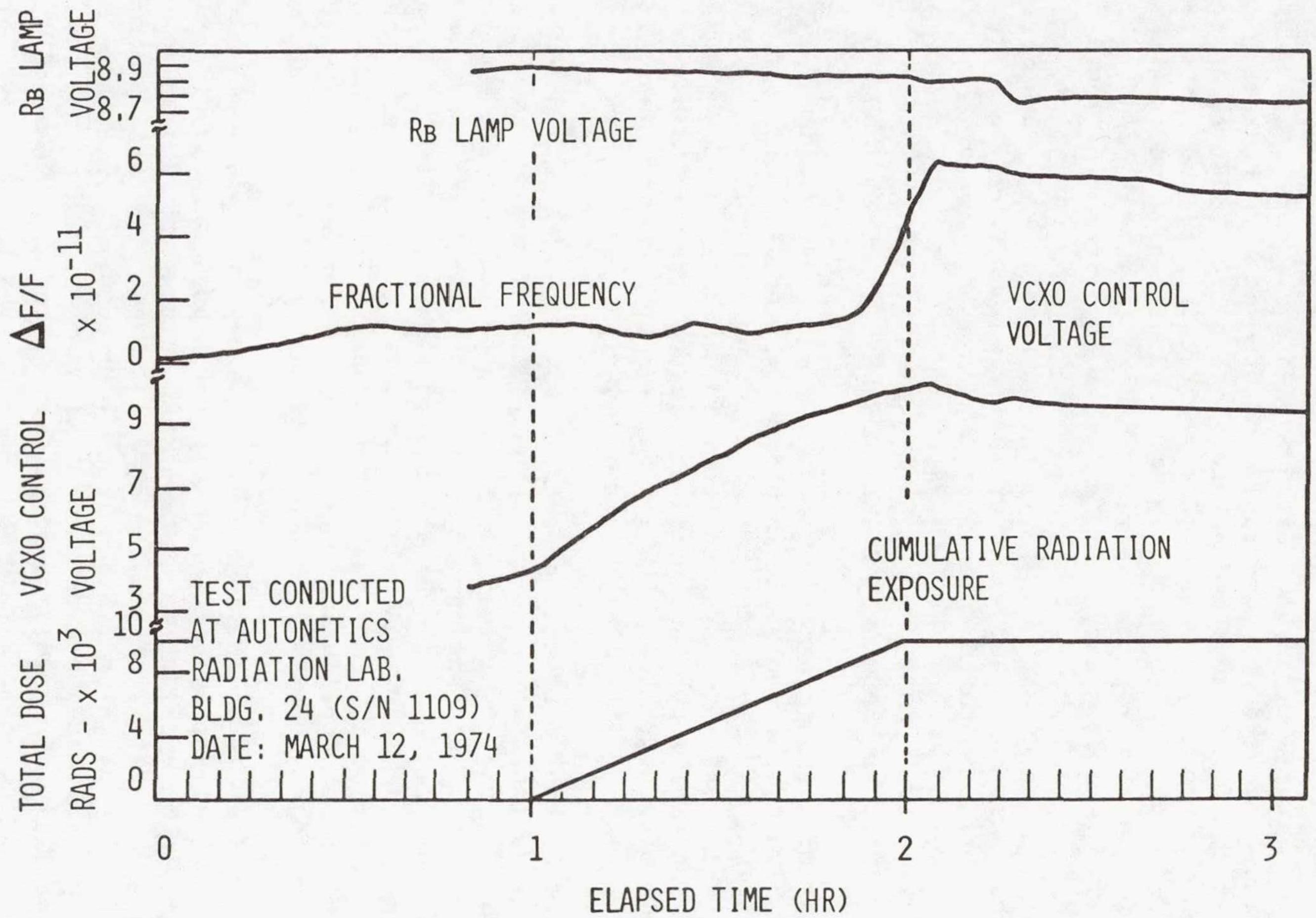


FIGURE 6. EFFECT OF 10^4 RADS (Co^{60}) ON AN OPERATING,

UNMODIFIED, SMALL COMMERCIAL RUBIDIUM STANDARD



This was not surprising, however, since the VCXO crystal was not designed or selected to withstand radiation effects.

In summary: The physics package of a small commercial rubidium frequency standard was essentially unaffected by a radiation dose rate of 4×10^8 rads/sec, and a total dose of 10^4 rads, in independent tests. For the dose rate of 4×10^8 rads/sec, the accumulated phase error was < 1 nsec, and this occurred within 60 sec after the radiation burst. For the experiment where the total dose was 10^4 rads, there was a frequency shift of about 6 parts in 10^{11} due to the irradiation. However, since this was due to a change in the electronics rather than to any changes in the physics package, and since the electronics in this case were not radiation hardened, this frequency shift could be eliminated in a carefully designed device.

APPLICATIONS ESPECIALLY SUITED FOR SMALL, LOW-COST RB STDS

Most of the applications for rubidium frequency standards utilize one or both of the two basic techniques listed in Table 5, namely, timekeeping, usually in the sense of measuring precise time and time intervals over periods up to about 10 hours, and also the generation of spectrally pure and stable microwave frequencies having high signal-to-noise ratios using the method of frequency multiplication from high quality low frequency signals.

Most of the applications peculiar to rubidium, as opposed to cesium, utilize the small size and weight, the low power consumption, and the low cost that rubidium provides. By far the largest application for ribidium at the present time is the use of these standards for navigation purposes in light and medium aircraft. We discuss this in some detail below. A related application is the use of these standards for positioning and geodetic survey purposes. An example of a positioning application would be locating the correct position at which to place an offshore oil and gas drilling platform.

Another class of applications is the use of these devices for secure communications systems; i.e., for military communications systems. This is an application that is just getting started and of which we will see quite a bit in the comming years. The area of secure communications can be divided into two groups: The first is message modulation and synchronous demodulation which uses the timekeeping capability of rubidium devices. The second is the use of spread spectrum techniques such as frequency hopping and pseudo random noise phase modulation that require spectrally pure and stable microwave frequencies with low phase noise, which is the second technique listed above. Again, the small size and weight, low power consump-

TABLE 5

APPLICATIONS ESPECIALLY SUITED FOR

SMALL, LOW-COST RUBIDIUM FREQUENCY STANDARDS

BASIC TECHNIQUES:

- TIMEKEEPING (PTTI)
- GENERATION OF SPECTRALLY PURE & STABLE MICROWAVE FREQUENCIES
  (USING FREQUENCY MULTIPLICATION)

APPLICATIONS:

- NAVIGATION (SMALL-MEDIUM AIRCRAFT)
- POSITIONING & GEODETIC SURVEY
- SECURE COMMUNICATIONS
  MESSAGE MODULATION & SYNCHRONOUS DEMODULATION
  SPREAD SPECTRUM (E.G., FREQUENCY HOPPING; PRN PHASE MODULATION)
- DIGITAL NETWORK SYNCHRONIZATION & MULTIPLEXING
- FREQUENCY CONTROL & CALIBRATION
- TIMEKEEPING PER SE (CLOCKS)

tion, and low cost make rubidium more suitable for applications of this type which require portability, such as in moveable field stations and military aircraft.

A somewhat related application is the use of rubidium standards for the synchronization of digital networks. This includes civilian, as well as military uses. An example of this is the Datran commercial communication system which uses rubidium standards for timing purposes (R. L. Mitchell, "Survey of Timing/Synchronization of Operating Wideband Digital Communications Networks," Paper 11, Session IV, this conference (10th PTTI)) . The last two applications in Table 5 are not especially suited to rubidium, except inasmuch as cost is a factor. In any case, these two applications are two of the more conventional ones as regards atomic standards.

NAVIGATION APPLICATION--RADIO NAVIGATION (VLF - OMEGA)

In Table 6 we are talking about the use of rubidium frequency standards in VLF & Omega navigation systems. The users here are owners and operators of light-medium aircraft. This includes both Lear jets and helicopters. In this application, price is a very important consideration. These types of radio navigation systems are typically priced in the range of $40,000 to $50,000. By way of comparison, inertial navigation systems sell for more than $100,000 and up. It is worth noting that it is obviously impractical to use a cesium standard costing about $20,000 in a radio navigation system that sells for $40,000. For this reason, the small commercial rubidium standard is the clear choice for this application.

A conventional VLF-Omega navigation system, which does not use an atomic standard, uses the hyperbolic method of locating position. In this method, a minimum of 3 VLF and/or Omega stations is required. Sometimes, radio conditions are such that it is not possible to receive as many as three stations. In this case, the accuracy of the system is greatly degraded. Even if three stations can be received, it may not be possible to obtain an accurate position determination. This depends on the geometical positions of the stations relative to the aircraft and the signal-to-noise ratios of the received signals.

A VLF-Omega navigation system that uses a rubidium standard does not suffer from these disadvantages. The inclusion of the atomic standard in the plane's navigation system allows the rho-rho navigation method to be used instead of, or in addition to, the hyperbolic system. The main advantages of the rho-rho system are that it is simpler to implement and is more accurate under adverse conditions.

TABLE 6

NAVIGATION APPLICATION

RADIO NAVIGATION (VLF - OMEGA)

USERS: LIGHT - MEDIUM AIRCRAFT, INCLUDING HELICOPTERS

NO RUBIDIUM STANDARD: HYPERBOLIC, MINIMUM OF 3 STATIONS REQ'D

WITH RUBIDIUM STANDARD: RHO - RHO, ONLY 2 STATIONS REQ'D

NAVIGATION ACCURACY

$$\Delta X = C \cdot \Delta T$$

DISTANCE ERROR = 1 FT/NSEC x TIME ERROR

CLOCK OFFSET OF 4×10^{-10} GIVES:

$$\left.\begin{array}{l} \text{TIME ERROR} = 6\ \mu\text{SEC} \\ \text{DISTANCE ERROR} = 1\ \text{MILE} \end{array}\right\} \text{IN 4 HOURS}$$

COMPARE WITH INERTIAL NAVIGATION ERROR OF $\sim$4 MILES IN 4 HOURS!

In the rho-rho method, the distance to a radio navigation station having known position is determined by measuring the time T that it takes for the radio signal to travel from the radio station to the aircraft. The distance X from the station to the plane is then given by $X = C \cdot T$, where C is the speed of light. The distance of the aircraft from the radio station defines a line of position (or locus) that is a circle of radius X with its center at the station. If the distances from two such radio stations are known, then we will have two such circles, one centered on each radio station. The aircraft is then located at one of the two points of intersection of the two circles. In this method, the distance error, delta X, is related to the time error, delta T, by the equation shown in Table 6, where C is the speed of light. In this equation the time error, delta T, is the accumulated time error of the atomic clock since the aircraft left its point of origin (point of clock synchronization).

Even if the atomic clock has a large average frequency offset, the navigational accuracy is still quite good. For example, suppose the average frequency offset of the clock were as large as 4×10^{-10}. Then the accumulated time error over a 4 hour period would amount to approximately 6 microseconds, and this would give a distance error of only about 1 mile. It is interesting to compare this navigational accuracy with that attainable by inertial navigation. For inertial navigation, the error would typically be about 1 mile for every hour of flight time, or about 4 miles in 4 hours! The rho-rho method is therefore capable of greater navigational accuracy at considerably lower cost.

FUTURE IMPROVEMENTS IN SMALL RUBIDIUM STANDARDS

Table 7 shows some of the improvements that can be expected in rubidium frequency standards in the future. We can expect the size to decrease by about a factor of two from the present small rubidium size of 1 liter. This will be accompanied by a weight reduction of about 40 % and a power reduction of about a factor of 2. In addition, we can expect warmup times to decrease further, by about a factor of five for a room temperature ambient. At -55 $^{\circ}$C ambient, warmup times of less than 5 minutes should be easily possible.

The temperature sensitivity will be less by at least a factor of 4. At the same time, it should be possible to reduce the sensitivity to changes in barometric pressure by about an order of magnitude. As quantities increase and manufacturing techniques improve, the price will decrease at the same time. It is difficult to predict this with much accuracy, but a price decrease of approximately a factor of 2 is reasonable to expect.

TABLE 7

FUTURE IMPROVEMENTS IN SMALL RUBIDIUM STANDARDS

CHARACTERISTIC	PRESENT	FUTURE	IMPROVEMENT
SIZE (CU IN)	67	34	x 2
WEIGHT (LBS)	3	1.8	x 1.7
POWER CONSUMPTION (W)	13	7	x 2
WARMUP TIME (25 °C AMBIENT)	10 MIN TO $< 2 \times 10^{-10}$	< 2 MIN TO 5×10^{-10}	$\sim$ x 5
TEMPERATURE EFFECT (-55 °C AMB. TO + 71 °C B.P.)	4×10^{-10}	$\sim 1 \times 10^{-10}$	$\sim$ x 4
ATMOSPHERIC PRESSURE EFFECT (SEA LEVEL TO 40,000 FT)	8×10^{-11}	$\sim 1 \times 10^{-11}$	$\sim$ x 8
COST (K$)	4 - 6	2 - 5	$\sim$ x 2

To summarize: The main improvements will be in the areas of size, weight, power consumption, warmup time, and environmental sensitivity. Other characteristics will either also improve, or else remain about the same as they are now. This should result in a wider range of aplications and concomitant lower prices.

DISCUSSION FORUM: ATOMIC FREQUENCY STANDARDS

HYDROGEN

Harry Peters, Sigma Tau Corporation

The things that I would like to compare hydrogen to, today, are not NBS-5 or 6 or the latest basic standard or NRC absolute standards, which are in great array here now; but with respect to present and future commercial cesium and rubidium because this is what I think is missing as far as hydrogen devices are concerned as has been pointed out today: they need to be available before they are going to be useful.

All the hydrogen masers today either originate in government laboratories, are government built, lent, or supplied, or are carefully tended antiques. And there are many examples of people desperately using hydrogen masers today. But there are lots of data to substantiate the performance on operating characteristics, to document their performance, present and potential, and improvements that may occur.

I am going to show two viewgraphs now to illustrate a couple of additional points. Now this viewgraph is rather a rough one. I apologize for it. It illustrates the rubidium and cesium passive standard systems and hydrogen maser active system.

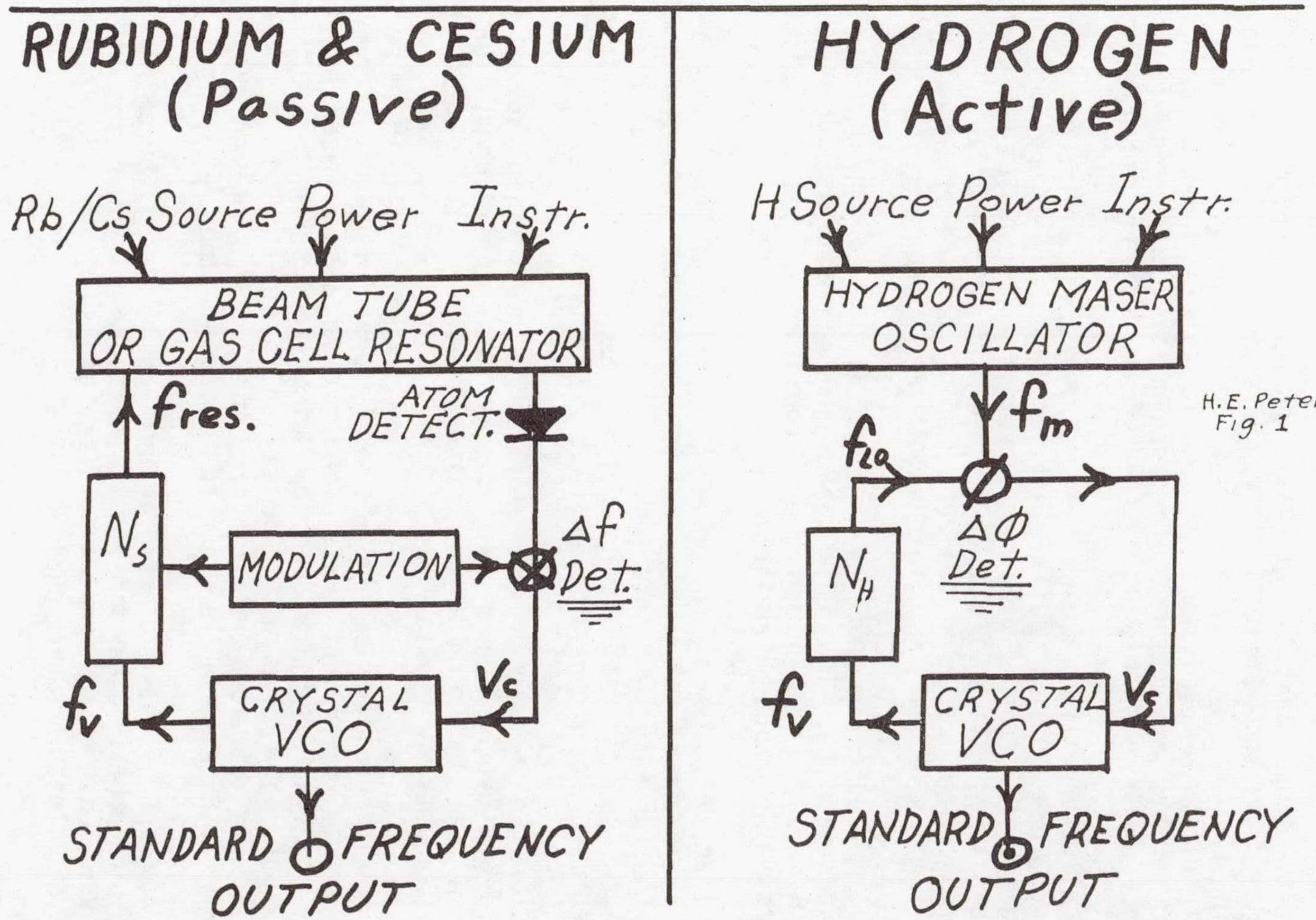
I am only showing them--not for a course on how beam tubes work and masers oscillate, but rather to illustrate the relative complexity of the systems. And for this purpose, both rubidium and cesium are well known to be resonant devices. But each of these devices needs a source of atoms. You have power input and instrumentation.

Then you have a crystal oscillator, which is multiplied up, if you are using a synthesizer system, to the resonant frequency. And you sweep the resonance by using a modulation frequency. You detect it synchronously and lock the crystal on.

Now let us look at the active hydrogen. You have a similar source, power input, instrumentation. You have an active maser oscillator, so it is a case of having a good low noise receiver to lock on to a coherent output. You have again the crystal oscillator and a number synthesizer to get the local oscillator. And you lock on the crystal using a VCO, and you have again the standard frequency output.

Now all of these electronic systems are really becoming very simple today, and electronics and instrumentation and power are of comparable magnitude, cost-wise, complexity-wise, and so forth. So they should not be large factors in future availability.

ATOMIC STANDARD SYSTEMS



H. E. Peters
Fig. 1

They are now because hydrogens have not been commercially developed. They have not gone through a large production run, and there has not been a lot of reliability work done. Most of them have been built by scientists in laboratories.

Let me carry on with the next viewgraph please. This viewgraph gives my own opinion, and I hope others, of the particular features of the hydrogen maser. I have just listed them. These are the only two viewgraphs that I will show.

First, it has the obvious stability characteristics, which for a long-term operating device are really outstanding and fundamental. Their reproducibility is exceptional. Now I intend this to illustrate it is better than any other device of the two types which I am comparing with. And basic accuracy: It is well known that hydrogen masers reproduce the cesium frequency to a factor of five or ten better than commercial cesiums, and in that sense they are a better basic standard.

This partly arises because of intrinsic reproducibility of the hydrogen frequency. If you look at the cost of the hydrogen maser per part in 10^{15}, you will find that it is several orders of magnitude less expensive than other standards where you need them.

The same is true of cost per year. The amount of hydrogen used today in a maser is trivial. The pumps last for decades, and you don't necessarily have to take them apart and replace the insides if the cesium becomes depleted or contaminates the tube.

They are simply active oscillators, and they can be made passive incidentally. There is work going on today in at least two laboratories, successful work with passive masers. But we all have our enthusiasms in this regard.

Reliability and longevity have been shown by papers which have been published. We don't have as much information as we would like because they aren't available in great numbers and for the reasons I mentioned. They are technically well developed.

The last point, I think, is quite obvious. They are not commercially available, so it is sort of unfair of me to compare hydrogen masers with commercially available rubidium and cesium. However, I hope perhaps you can say I am comparing them as future ones.

Let me go right on into applications. I am not going to go into detail. I think that a lot of this detail has been gone over today, and everyone knows where you can use parts in 10^{14}, 10^{15}, etc. and perhaps where it isn't needed.

FEATURES OF HYDROGEN MASER

- STABILITY
- REPRODUCIBILITY
- BASIC ACCURACY
- COST Per PP/10^{15}
- COST Per YEAR (CS Ens.)
- Simple Active Oscillator (Or Passive)
- Reliability
- Longevity
- Technically Well Developed

H.E. Peters
Fig. 2

NOT COMMERCIALLY AVAILABLE!

We need them of all places, obviously, where rubidium or cesium are not less costly, or are not adequate for the application. For example, a very important application is that a given hydrogen maser--you might want two actually for redundancy--can replace a very large cesium ensemble in principle in basic timekeeping systems. Of course, you have military and NASA ground stations, navigation and communication stations, time and frequency calibration labs, astronomy, VLBI, geodessy, and other scientific and military uses.

Next I would just mention a couple of words about future performance. We have at least three laboratories that are either at or pushing into the parts in 10^{16} region for certain averaging times. And I would like to point out that some of this work is being done with masers which have aluminum cavities, and other masers have dielectric cavities with a similar performance.

Now, with regard to the types that I have been using, I am not really an advocate of aluminum basically; however, it is not clear that the instabilities are due to temperature in many cases. I would point out that in the future I expect in the aluminum type, by using subsidiary dielectric materials, that you can easily get a factor of 50 in principle and you get down into the part in 10^{16} region in stability just due to the further lack of pulling, due to cavity pulling.

As to size, there are new approaches available. They have to be proven in the laboratory, and most of you know about these attempts. But the best standard is, in principle, the largest standard, and that is all I will say about that.

The future availability depends entirely, of course, upon getting well-known state of the art into private industry. You cannot sell, distribute, support, or go around in the field and have people to maintain or have wide usefulness in a hydrogen maser standard if they are built in specific scientific laboratories and don't go through the disciplines of the production cycle.

Just one more point I should make. There is one place in the world today with significant private investment to achieve both active and passive hydrogen maser frequency standards. It is not in the United States. It is a well-known company in Switzerland, Ebauches. And if there are any further questions regarding the work there, I would refer you to Dr. Busca, who is present today.

I think that we will see hydrogen masers in the future in the United States. I personally think that, unless we can get some way to stimulate availability of commercial standards in the United States, there is a high probability that we will have Volkswagon hydrogen masers, Le Car hydrogen masers, Toyota hydrogen masers, etc. It doesn't appear that we are going to have any Ford, General Motors, or Hewlett-Packard hydrogen masers in the near future.

Page intentionally left blank

OPEN DISCUSSION: FORUM ON ATOMIC FREQUENCY STANDARDS

DR. HELMUT HELLWIG, National Bureau of Standards:

I would like to thank all of the speakers. Before we open the discussion, Raymond (Besson), could you join us. You know, we have sort of competing standards concepts, and that is the purpose of this panel.

Let me first ask simply, does anyone here agree, disagree, or have comments on what anyone else has said? Raymond, you are first.

DR. RAYMOND BESSON, E.N.S.C.M., Besancon, France:

Immediately I have a comment. I will say to Dr. English that I perfectly agree with his positive statement about the rubidium. And I really enjoyed the talk and the work in this area. But I do not agree with his negative statement about quartz.

You say you do not have available data about short-term stability. Well, if you don't, measure it. You should also take into account the last results, at least the last commercially available results. You say I don't have the numbers. You pointed out....

DR. THOMAS ENGLISH, Efratom California:

Why don't I just put it on the viewgraph. It is right here.

DR. BESSON:

Okay. On short-term stability, I would like to see some figure, whatever it is. Okay. Long-term drift, I would say I agree with the commercially available number.

MR. ERNST JECHART, Efratom California:

It is important. It means only commercially available.

DR. BESSON:

Yes. Okay. But you know I simply would not like to make a too partial point. I would like simply to see some figures for the short-term stability. Warmup time can be discussed. But, for instance, for acceleration sensitivity, I just stepped out here during the session and saw companies just giving some sensitivity data which is not 8×10^{-10}/g. So I think I would like some of those numbers, you see, to take into account the last data.

Also about the compensation, Don Emmons pointed out some results and compensation devices that are also available in France. It has been three years since Valdois did his work. And really I don't believe we come out with this terrible acceleration sensitivity right now, at least in France.

DR. ENGLISH:

This was based on commercial quartz oscillators that are available right now and represent, I felt, the state-of-the-art parameters. Now I don't promise you that I have the exact correct value in every case. I really honestly tried to give good values here. Now if the acceleration sensitivity is too high, I would appreciate it if you would correct me.

But these are commercial units. I am not talking about the state of the art for quartz in laboratories.

DR. BESSON:

Oh, naturally, but you should, I think, give a number for short-term stability because this is available on commercial units.

DR. ENGLISH:

Well, I did mention that usually the best I knew was parts in 10^{13} when it was specified. Usually it is not specified over 100 seconds. See, I have over here short-term stability, minutes to hours. I know it is usually specified out to 100 seconds. Beyond that I don't know.

DR. BESSON:

You know the point--it is always very difficult to make such comparisons because it can always be discussed. It is, rather, a feeling. Negative statements always lead people to drop some effort, and I believe, like I said in my talk, that there are many routes available. Don't make the people throw away the quartz oscillators.

DR. ENGLISH:

Well, I don't think we are going to throw away quartz oscillators because they obviously have advantages. But this was supposed to be a parochial presentation.

MR. JECHART:

Tom, I would also like to say that it is not fair to compare a commercial rubidium unit with a crystal that is in a laboratory. If we reverse this we can make a much better statement for rubidium. What we have done is....

DR. BESSON: I perfectly understand, but....

MR. JECHART:

Yes, but you said it was a negative statement. I don't believe....

DR. BESSON: Negative, yes.

MR. JECHART:

It is not negative. It is what you call negative. Please tell me one company that produces crystals on the market with a better value than 8 x 10^{-10}/g? One company? I don't know of one with a better value because we said commercially available.

DR. HELLWIG:

Maybe we should reduce the discussion at this point from what could be done in the future to what is available now.

DR. C. C. COSTAIN, National Research Council of Canada:

I just want to take one crack at Harry Peters here. We promised each other that. And this again is not connected with commercially available units, but with bringing things out into the open. Since I came into this business six years ago, Dr. Guinot and others have been pleading to have hydrogen maser clock measurements reported to the Bureau International de l'Heure (BIH) so that some evaluation against the international standards can be made. I would hope--we haven't seen it yet--that next year, three or four masers will have made ten months of reports (because they always come back 60 days late) and that at the next PTTI, we can have results on these masers that have been reported to the BIH so that some evaluation can be looked at.

MR. HARRY PETERS, Sigma Tau Corp.:

That is terrific, yes. I would like to see that also. However, the people that have been making hydrogen masers have been very few. And the evaluation and distribution of these standards to organizations makes it almost impossible.

Actually, I am not in a position to help with such a thing. I think there is some discussion of the possibility of Goddard masers being compared with TAI and so forth. Now, as to whether hydrogen masers have been observed, they have been observed for months, eight months or up to nine months, and for three months at a time in comparison with cesium ensembles, and have done remarkably well.

It is very difficult to confirm whether a great number of cesium clocks have been stable or hydrogen masers have been stable. So there is no doubt in my mind that there are a lot of data, both maser to maser and also internal data. It is done in our laboratory, of course, and it would be preferable to have everybody in the world look over the data. However, I believe in it, and they are as stable as theory predicts.

There is no opportunity to do what you say until we have several units which have gone through the discipline of being produced

and have the support, and other people can do it. Because the people who develop these, make them, invent them, cannot possibly take up the whole job of your organization, and compare them with TAI. It is an impossible burden, and I hope others will do that when and if they ever get masers.

Now, could I continue on to comment since Dr. Costain has opened the door.

MR. PETERS:

I want to thank you particularly for predicting presently unobserved effects in hydrogen masers due to magnetic effects. Actually, our holes get smaller as the size of the shields get smaller, and they are smaller than some have told about in magnetic shielding calculations and some of the later designs of masers, such as a spacecraft maser and another design which has been proposed. They are only about 0.4 inch.

Typically, these designs have been approached rather carefully, and we are already down to inches for bulb state selector distances, for example.

DR. COSTAIN:

It is not the holes. It is the material I am concerned about.

MR. PETERS:

No, no. I am going to relate to that in just a moment. As a matter of fact, you can easily evaluate the inhomogeneity shift in hydrogen masers through several well-known techniques. They have been published, and the effects do not create either an inaccuracy or a resetability problem.

I don't anticipate that we are going to make hydrogen masers an order of magnitude smaller than they are now, so it is not going to be an order of magnitude change.

I would like to make one last comment since I still have the opportunity. I wish you could have put your enthusiasm and your talents and your opinions originally into hydrogen because I think we would be much further ahead in hydrogen technology than we are today.

DR. COSTAIN:

We have a couple of the ancient masers that you referred to in Ottawa, vintage 1965.

MR. PETERS:

I have pictures of them.

DR. COSTAIN:

When the cesium program is finished, we fully intend to see if we can do better in hydrogen masers.

DR. HELLWIG:

Maybe I should make a comment here and conclude this part of the discussion. It is, I think, a perennial battle within every lab. I saw it go on at NRC, PTB, NBS, and, I think, in a derived way, at some, say, non-standards labs as well. You have certain tasks you have to do and that causes you to order priorities. And you are not always doing what from a purely technical viewpoint is the best choice. I think that goes for almost any decision we are making.

Let me change the topic slightly. It is very close to what we have started on. In many peoples' minds, again back to the systems designer, the user, etc., there is always a ranking of standards. Sometimes the ranking goes hydrogen, cesium, rubidium, crystal. Sometimes the ranking goes in reverse. It depends upon your requirements and how you look at them.

We, I think, have not fully addressed that yet in the formal presentations or in this discussion. Where are the actual niches for present day hydrogen, rubidium, cesium, crystal? And where are the potential niches? I'll ask Harry first.

Hydrogen definitely produces the best numbers. There is no question in terms of stability. What in your experience are the present customers, and what are you seeing as customers, either based totally on this exceptional stability performance of hydrogen or maybe on other qualities of hydrogen as you see them? In a nutshell, who needs hydrogen, now and in the future?

MR. PETERS:

I think that the answer lies in the fact that you don't absolutely need it, but it would be more economical to use it, and you would have a better system if you used it, and it would take fewer computers and fewer people if you used it. You would have better navigation. Maybe you don't quite need it, but it has many practical advantages.

I think the users are obviously time and frequency organizations such as your own, and also international timekeeping organizations where it has been cited by others that it would be desirable to have a comparison standard for the present cesium ensembles. Perhaps it is more clear now since we have got better accuracy than we had before. But actually, absolute accuracy is what determines our long-term knowledge of our frequency drift of any group of standards. It is not an ensemble of commercial devices which are not in themselves absolute in the sense that they can be evaluated.

So we have that application. I think that for certain navigation systems, Loran-C, Omega, they may develop needs as the capabilities of timing and frequency and communication systems evolve into the future.

Certainly, many applications of which I can't possibly be aware, military, mobile, and so forth, are possibilities. All of the NASA tracking stations, many of them, particularly DSN.

Incidentally, I think 75% of the hydrogen masers being used today, such as our old H-10's and many of the laboratory devices such as those Dr. Costain referred to, should be replaced with something that is a little more modern. And I am sure he agrees with me. I think I could go on. I think many university labs....

DR. HELLWIG:

Do you have a rough guess as to how many hydrogen masers are not only in existence but actually being used?

MR. PETERS:

I'm sure that if we counted carefully and not only in the Western countries, we would find 50 or 100 of them, something on that order. And there are more in the Eastern countries because they started out with more enthusiasm in this direction, I believe, than we did. At least, in Russia, their basic standards were originally hydrogen, and now they have come over to cesium.

Then there is geodesy and astronomy, and VLBI of course. Any system that needs phase coherence, where you have to resolve to 10^{10} cycles, one radian or a small fraction of a radian, really needs this type of standard. And if you can maintain this over a day's time, of course, and over long periods of time, you don't need to resynchronize.

I'm sorry. I've gotten away from customers into performance again. Did I answer....

DR. HELLWIG:

Yes, you really answered that question.

MR. JECHART:

What do you think is the price if it is available for commercial use?

MR. PETERS:

I think the electronics will basically be the same price as the electronics for your rubidium cell, almost.

MR. JECHART: Really?

MR. PETERS:

We have synthesizers on one board; they are operating and work beautifully. You have shown all your beautiful electronics. There is so much large scale integration. Temperature controls are a couple of very inexpensive IC's, for example. Power supplies are commercially available. They have just been put together by high-priced scientists in the past. Basically, the electronics part is not expensive. The large parts are made of aluminum, and if they are made in production, they will come down by a factor of who knows--two, three, five, ten, depending on how many. We are never going to make them like automobiles, of course.

MR. JECHART:

Do I understand you correctly, that you can make it much cheaper than cesiums?

MR. PETERS:

I think we can make the price comparable. I think we would probably put into the hydrogen things that cesium doesn't have because it (hydrogen) is inherently more reproducible and also has higher resolution on the C field. And you would want to, with the short-term stability, have much greater resolution on the synthesizer.

So we have put in a couple of little things that really make it this much more useful than the cesium would be because of its (cesium's) high-shot noise.

I think you can bring it down significantly in price, but with the fluctuation of the dollar and so forth, I don't want to say what it is going to be.

DR. HELLWIG:

I wanted to ask the same question about rubidium. Who needs rubidium now and who will need it in the future? But I think Tom English answered that question to a large degree, so I will modify it. Please answer this question with a special twist of thinking of either simpler or higher performing rubidiums, which I think Tom did not really address. So sub-question "a," is: Can rubidium be so much improved that it really competes on the level of present day cesium, maybe even hydrogen? Then, I think, the answers would be the same as those given by Harry for hydrogen, and you would have direct competition between the standards.

MR. JECHART:

I would say it is not a simple question.

DR. HELLWIG:

Sub-question "b" is: Could you even further simplify the rubidium to the degree that applications open up where, as I tried to point out this morning, at the present time, you have difficulties in coming up with standards at all. What is your answer?

MR. JECHART:

The answer to the first question is yes. We have experimental data that were not discussed here. And the data show that we can go close to cesium. This is what I believe, what my experience shows.

DR. HELLWIG:

At no great increase in complexity?

MR. JECHART:

No. The answer to the other question is yes because with modern electronics, you can make it much cheaper and smaller, as you can do with everything. But because the rubidium physics package alone is already so small, it makes sense to make an electronics package much smaller. Tom showed, on the last graph, what we believe is possible in this area.

DR. ENGLISH: Factor of two or so.

MR. JECHART:

Yes, and this is really what your second question was, I think; and of course much cheaper.

DR. HELLWIG:

Let me repeat what I think was an important point with regard to rubidium. As contrasted to hydrogen and cesium, the electronics is the bulk of the size at this point. So that gives a totally different attack angle for the designer of the clock.

MR. JECHART: Yes.

DR. HELLWIG:

Cecil, the same question. Where is cesium used, where will it be used, and where do you think fundamental improvements could be made?

DR. COSTAIN:

Well, I think in any of the standards that any step in accuracy that can be achieved is immediately saleable. I don't think there is any question there, and it is just a question of commercial viability.

In cesium, certainly, there are going to be limitations. I think it is almost inevitable if you reduce the size, you are going to reduce the accuracy.

We thought we'd see if we could put one in a watch one day, but we didn't get very far. The magnets might go in a watch, but the microwave source is a little difficult. But perhaps you could come down in size and not sacrifice the accuracy totally.

What we have done in the lab is perhaps an experiment in practicality. We don't know yet. We will hopefully know next year if you can match, in one device, what you can do by averaging 100 commercial units. In that case, there might be requirements in ground navigation, or you might say absolute indexing, although I think it is more fun doing the indexing by satellites, and I hope we can discuss that on Thursday.

DR. HELLWIG:

Thanks, and finally, Dr. Besson, you get your chance. Crystals. You should not talk about commercial crystal but about some fancy devices of the future. Will they wipe out atomic standards? If so, how?

DR. BESSON:

First, I would still go back to normal units and point out that quartz crystals are still the work horse in frequency and time measurement systems, since almost any device has a good quartz crystal. That would be my first point. I should have said it sooner.

DR. HELLWIG:

Excuse me. That means if the world is populated exclusively by atomic standards, which I think is nonsense, that there would be at least an equal number of crystals?

DR. BESSON:

Yes. That is exactly what I mean.

MR. JECHART:

I don't know if atomic standards always need a crystal.

DR. BESSON:

Yes. So that would be my first point. Second, it is always very difficult to speak about the future because one has a tendency to be optimistic. But I believe in some qualities of the quartz oscillator. It is a low-cost unit. It is low volume and can be operated with low power. At that point, I very strongly believe that the g sensitivity will remain a problem for quartz units.

It all depends to what extent. I would say that down to some parts in 10^{-11}/g, I don't see any problem right now. And it is a very important point because some years ago, this was still a problem which was not solved at all. And it is a very difficult one because you do need theoretical stuff of high level. And you do need to realize it experimentally; it is not enough to make nice calculations.

So I can tell that we solved the question in France. We made our first low-g-sensitivity resonator, very low-g, one year ago. This is done, and you can reproduce it very well. Of course, if you would like some g sensitivity down to 10^{-12}, that is another step.

So the next question, "Will it be a huge market or not?" I don't believe I can answer. I am working in a research lab. But I believe that those numbers could certainly bring some customers.

MR. JECHART:

I feel what is important here is not what you say you can do, but if you do it. Of course, I am sure you can, but if you compare later on what is important for a customer, it is really the price also.

DR. BESSON:

Okay, I already said that the price of the unit we are evaluating right now should range, at least for as far as I can see, from a factor of 1.1 to 1.7 of the price of regular units. But we don't believe there will be a large increase in price.

DR. HELLWIG:

Excuse me. Less than a factor of two increase in the complexity of manufacturing, right? As compared to normal?

DR. BESSON:

Yes. And also you do have to know some features--I am speaking now about the new crystal--which are very nice, like frequency adjustment, which allows you actually to get rid of, to a certain extent, the series capacitor. It is too soon to say things now, but I believe that there is some kind of hope for the very near future.

DR. HELLWIG: Questions?

DR. VIG:

As a quartz crystal man, I was very happy to hear Raymond Besson defend quartz crystals. And I think also he is being much too modest as to what the future holds for quartz.

Those numbers that you mentioned before for commercially available crystals were all numbers for singly-rotated, either AT or BT

cut crystals. For those of you who are atomic and molecular frequency standards people, there is a revolution taking place in quartz crystals in that the double-rotated cuts, in particular the SC cut, are known to be much less sensitive to stresses than the AT cut. And this has produced improvements in short-term stability, which was reported already at the Frequency Control Symposium last year. And also this morning, it was mentioned that stabilities of 10^{-14} per 100-second averaging time, I believe, were achieved already.

I would also like to point out that the crystals that were mentioned, as far as being commercially available, are usually manufactured using technologies which are 10 to 20 years old. And there are some technologies that have come along in crystal fabrication, such as in cleaning and packaging, which again will probably produce orders of magnitude improvements in long-term stability.

I would also like to mention that even though the long-term stability mentioned in that chart and in the HP catalog, for instance, is like 4 or 5 parts in 10^{10} per day, it was mentioned by Jack Kusters at the last symposium, for instance, that the actual crystals are aging in parts in 10^{12} per day. Even though they are not going to guarantee that in their catalog, the actual units do age parts in 10^{12} per day, and that is still for a singly-rotated crystal, where most likely the dominant aging mechanism is in the stress relief.

If you can use the most modern fabrication techniques for eliminating contamination plus use the SC cut for eliminating stress effects, there is every reason to believe that orders of magnitude improvements in long-term stability will result.

DR. HELLWIG: Is there any comment from the panel?

DR. BESSON:

Well, as a fact right now, I think we know very well a way of making SC crystals; for instance, their g sensitivity would be less than 5×10^{-11}. And that is not one crystal. That is a lot of them. And then you can compensate if you don't like this 5×10^{-11}. And I do believe that more improvement can be made.

I think that the quartz business is really at a turning point. I pointed that out in my talk. For 20 years, for some reason, there were some kind of asymptotic performance that caused people to maybe be less interested in quartz. But I believe this is going to change.

DR. HELLWIG:

I think I know the reason for the 20 years. I can quote correctly, I believe, Don Hammond of Hewlett-Packard, telling me that the

advent of atomic standards stunted further scientific and advanced engineering development of crystals. That was 20 years ago.

DR. BESSON:

I really think that this is true, for instance, when I think about the techniques that John Vig has developed right now. There is an incredible amount of work that is being done now, and we are just ready to gather the benefits from that. And there is this technique of John's where the packaging of crystals can really bring much in the result and stability and drift per day. And one day I think we will put all those techniques together. We would like to use more of your new packaging or chemical etching and things like this. I believe it is time to do it, and I believe that the results will be surprising. So I perfectly agree with John.

DR. KAHAN:

I would like to argue that same area that John Vig and Dr. Besson are arguing. Personally, I see rubidium becoming obsolete very shortly.

We have to understand that the standards are competing against a moving target. Now what you have heard today or at the previous symposium is tremendous development both in quartz oscillators and resonators, and performance. In terms of cesium standards, what we have heard the last few years is not so much development in terms of performance, but development in terms of operational parameters, cost, size, and system limitations.

From what we have heard in Tom English's first paper, which is a complete contradiction to his second paper, for example, is that in one case he is concentrating on taking on the cesium market. And yet his last slide in the second paper is the possible improvements that can be done in rubidium, which doesn't mention long-term stability at all. What he mentions very properly is weight, power, and warm-up in terms of airborne navigation. And in that respect, I think that market will disappear as soon as the quartz oscillators become available.

I think that in that respect, cesium is moving ahead and quartz is moving ahead, but I haven't heard anything either today or at the Frequency Control Symposium which makes me believe that rubidium will soon be a valuable standard a few years from now.

MR. PETERS: Could I give a word for rubidium?

DR. HELLWIG: Yes, I need a word for rubidium right now.

MR. PETERS:

It is probably not well known, but some of my first work in graduate school was with rubidium cells and spin exchange. So you are forewarned.

Actually, it seems to me that quartz crystals are very good when you need them. They are never going to disappear. It seems to me also that for the very long term, atomic standards are always going to be better than mechanical standards.

MR. JECHART:

You say that right. It is a mechanical resonator, and I feel that this is a limitation you don't really have in an atomic standard.

I'm not sure you said the aging was much better. So of course, you can find the same thing in rubidium. We have rubidium standards that don't age, but the important thing is what you can tell a customer you can guarantee--and here I feel is the limitation.

What I heard regarding crystals is fine because you now use a new technique and you can make improvements. But first, what is the price? Is it economical compared to rubidium, and what is the advantage? I feel we here in this room should really discuss this fairly in a technical way. And I think this is possible.

Back to the crystal: Of course, for me, the limitation is the mechanical resonator. This is what I believe.

DR. BESSON:

I would not like to make any negative statement about rubidium standards. The first question is, what do we need and what do we want? And we ought to know what is available.

I really find once more that, you know, one should not make any negative statement. I think it would stop progress for a while. The same thing is true if you say, "Well, bulk devices, that is okay, but you have got saw devices now, and bulk devices are going to die."

MR. JECHART: Look, what I am saying is.....

DR. BESSON: You may be wrong.

MR. JECHART:

I am sure it can be much better with the new technique. On the other hand, look at the rubidium field. It was also the same way. Nobody really spent money in this field, and nobody worked on this (basic advances). And I am sure the rubidium situation could be the same as you are doing now with crystals. You can always make

improvements. I feel it (rubidium) is not limited. This is what I would like to say.

DR. HELLWIG:

Well, we have come back to the old point, I think, of having a need for something, and only then will it be produced in adequate numbers with adequate performance characteristics. It will be produced if it can fulfill a need which the other devices cannot fulfill under the same conditions of, say, size, power, performance, warm-up, cost, and so on.

I think the question, put differently, is that some people here think that crystals can assume certain characteristics which were not possible before and were only available at a reasonable price from rubidium standards. But I think it is probably too simplistic to think that that makes rubidium unnecessary, for reasons which were stated already. And I think the same is true for cesium and hydrogen......

MR. JECHART: Yes, every one has a place.

DR. HELLWIG: Andy Chi has a comment.

MR. ANDREW CHI, NASA Goddard Space Flight Center:

I would like to first commend the panel members. That is, the presentations were parochial and also very interesting. However, one should realize the different types of atomic standards, including crystal oscillators as oscillators, which are competing, are not really the same. They have a certain amount of common characteristics. It is hard to believe that one can be substituted for another if a sophisticated user has genuine need for a particular type of standard with particular specifications. Most likely, he would be dictated to use one type, perhaps two. But they are not all the same, which is the impression that is given.

The other point I would like to make is the fact that in the applications, it is very hard to identify a particular application for a particular type of standard unless one can specify the requirement needed for the application.

One can conjecture and guess. This almost brings back to mind when cesium standards were developed. The estimated number of potential sales was 50. And you can see that now the number of sales of cesium is more than 50. And of course the same would apply to hydrogen masers.

Now rubidium by itself has its own use. In short, I am not sure it would be replaced by a crystal ocillator. In the same way, crystal oscillators will never be replaced by atomic oscillators.

DR. WINKLER:

I think Mr. Chi has stated a very important point. I think one can make a case that, at the moment, about ten hydrogen masers are being built and sold per year, about 100 cesium standards, about 1,000 rubidium standards, about 10,000 quartz crystal oscillators of the quality which goes into a frequency counter or similar type of instrument, about 100,000 low-quality quartz crystals per year, and possibly a million going into the quartz crystal watch industry.

So, why are these standards or these oscillators being used in these almost decades of orders or quantities? I think there is a very good reason for that, and that is that each one of them has certain performances which attract a certain clientele.

I am very pleased to hear that, in all of these devices, very interesting and most promising progress is being achieved. Now if I may add a few other comments, going back to some of the other things that have been said before, I think there is one misconception in regard to the lifetime of the high-performance cesium beam tube.

It is true that there has been a higher failure rate of those than the regular ones. But I believe that it cannot be an intrinsic great difference in lifetime because we have several of those standards performing very well after five and six years. And so I think we have to distinguish manufacturing and quality control problems which seem to have existed from intrinsic problems because the first ones eventually get straightened out. The second ones require a different design or engineering approach.

Regarding the magnetic comments of Dr. Costain, I really wonder whether you are talking about final limitations coming from the shields or coming from the random remagnetizations, random magnetic reorientations of the total material in your transition regions.

DR. COSTAIN:

Yes, the total material, of which I think the shields are the most important. But speaking really not of the shielding but of the shields themselves, we have found, by monitoring along the length of the tube, really unexpected changes in the field.

DR. WINKLER:

But then I think we have arrived at the paradox that, if I assume your numbers are parts in 10^{15} for the long standard and parts in 10^{14} for a two-meter beam, a normal so-called commercial standard ought to be not better than parts in 10^{12}. And this is a little bit difficult to conceive.

I also think that these effects are not borne out by the experience with hydrogen masers, as I think Harry Peters has already hinted at. I feel there may be something else or a different kind of manufacturing.

Incidentally, that brings me to that question of terminology. Wouldn't it be better to not distinguish oscillators or clocks by the particular way in which their development has been financed, but by the difference in manufacturing. Here we talk about a laboratory device or industry-produced device, which I think is a more significant difference than to talk about commercial and .... I don't know what you would call the other one. I think the distinction is one that has to do with which way do you manufacture them because there are some individually built devices in an industrial environment that have performed exceedingly well. In fact, one-half of NBS-4 actually originated this way (by industry-production). Isn't that true, Dr. Hellwig?

DR. HELLWIG: (Nods affirmatively)

DR. WINKLER:

So I don't know. Regarding the contribution to the BIH, Dr. Costain does not read carefully enough the bulletins of the BIH because there is a hydrogen maser contributing to the BIH. And, believe it or not, it is at the Naval Observatory--with some interruptions, I must agree.

Regarding the phase noise curve that has been shown by you, Dr. English, isn't it true that all phase noise contributions coming from about the breaking point of your servo loop really are the crystal's and have no bearing whatsoever on whether this is a rubidium reference or cesium reference or hydrogen reference? Am I correct?

DR. ENGLISH:

Could I make a comment on it? You are correct. I think the problem is to do it in such a way that you don't add to the cost of the unit. And if you look at the rubidium, the cost that it adds to achieve that phase noise is pretty minimal.

DR. WINKLER:

Yes, but the industrially-produced cesium beam standards have a very good crystal oscillator also. The trouble is you don't see it because it is shielded so much by buffer stages for the purpose of avoiding external interference signals going into the transition region; and this would indicate to me that your rubidium does not have this kind of buffering.

MR. PETERS:

Could I just comment on one thing? I just wanted to mention there is one little difference in crystals in a hydrogen maser rather than the passive device. A passive device that gets a glitch in phase on the crystal will not recover because it is a passive resonance. A hydrogen maser has an active lock on a crystal and it will recover

in an infinitesimal amount of time any phase lost. So there will be a difference in the statistics of your results.

MR. JECHART:

Dr. Winkler, about your comments on phase noise in cesium devices. Perhaps I do not understand what you are saying because the servo loop time constant is much larger for cesium than for rubidium. For example, our rubidium has a loop time constant of 100 milliseconds. This means that for times shorter than 100 milliseconds the stability is due to the crystal, and for longer times it is due to the rubidium. And in cesium I know that the loop time constant is selectable, either 1 second or 50. So even if you use a very good crystal, you still have an influence from the cesium tube, and this is the reason why it is . . . .

DR. WINKLER:

Exactly, but I was talking in the frequency domain and talking about frequency offsets larger than 10 Hz or 100 Hz.

MR. JECHART: Yes, this is the crystal....of course.

DR. COSTAIN:

Just one final word on the magnetic domains. I think in the industrially-produced standards, you certainly do have effects that are a part in 10^{12} or larger if they are not very carefully degaussed. And I was pointing out that this scale factor is a cubic one and that you can run into trouble awfully quickly if you don't expect it. You have got to be much more careful in the degaussing and, in fact, in the construction of your shields, which might be easier in one type of device than another, and determine if a weld or riveting or seals might seriously influence the device once it gets small. If you are a foot away, it doesn't matter.

DR. HELLWIG:

Yes, I was just reporting one result on NBS-4, which is half a meter long, a little shorter than your present CsVI, A,B,&C. If I remember right, our magnetic field limitations are below the 10^{-14} level in that device. We know that from measured data. Are there any more questions, comments?

MR. SAM WARD, Jet Propulsion Lab:

I have two questions and one comment. In listening to the relative performance there, it appears to me that with hydrogen masers and cesium, it's think big, and for rubidium and crystal, small. So, based on relative performance, per unit volume, the rubidium is the clear winner.

Now for the question. Has anyone for the crystals considered using the separate cuts of crystals in a mixed device, using a crystal that is most favorable for the long-term performance to slave to a device that has the good short-term?

And the second question, addressed to Harry Peters, that those of us using hydrogen masers in a widely dispersed net have a great need for better accuracy because the cost of establishing synchronization is really a heavy burden.

DR. BESSON:

Well, I think that the answer to the question for quartz crystal is that this possibility has been demonstrated. That is at least what we can say. For instance, by just using the crystal that was driven at--that was some years ago--low power for long-term stability and run one at higher power for short-term. But now we do have some better things to do with two crystals. And I think it has been demonstrated; you simply have to make it in a very smart way.

DR. HELLWIG: Harry, do you have a comment?

MR. PETERS:

Well, I think certainly hydrogen masers at deep space networks and so forth, systems like that, will make a contribution to the basic reproducibility of the frequency at the stations. Of course, we still have to establish epoch, but you won't have to resynchronize so often. I thank you for your comment on the need for hydrogen masers.

DR. HELLWIG:

Maybe I should insert a comment here. There is a need for, as I call it, syntonization. It means equal frequency, and we normally assure equal frequency by electromagnetic signals or by portable clocks, which are really time difference measurements over time intervals. And I think we are coming to grips with the possibility of establishing from scratch a frequency with very high accuracy. And I think the hydrogen maser may be the first choice, at the moment at least, to carry frequency around. Not with an operating device: turn it off, ship it, turn it on, and it is within certain narrow limits, I think. What would you say the limits may be? Turning it off and then on again; that is, reproducibility?

MR. PETERS:

Oh, I think we have basic intrinsic reproducibility in principle. It depends on whether you have an autotuned device or one without all the sophistication you might want to put into it.

DR. HELLWIG: What is technically feasible? What is the number?

MR. PETERS:

Oh, I think our hydrogen masers will reproduce in the range of at least parts in 10^{13}.

DR. HELLWIG: That is very conservative, I think.

MR. PETERS:

Well, I don't want to take any biased or unconservative stand. Thank you.

MR. WARD: What I meant was absolute accuracy.

MR. PETERS:

Oh, it depends upon whether we are talking about intrinsic reproducibility or reproducibility like we compare the frequency here and then we compare it after it is turned on. And that might be better than the one I gave.

DR. HELLWIG: Excuse me. How do you define absolute accuracy?

MR. WARD: In traceability, for instance, to A1.

DR. HELLWIG: You mean in terms of time or frequency?

MR. WARD: Frequency, syntonization.

DR. HELLWIG:

Okay, but you mean in reference to some established standard, in which case reproducibility would be sufficient?

MR. WARD:

Yes. For instance, in our net we have to maintain so many parts accuracy. And the length of time over which we have to maintain this is longer than we can keep a single unit working. And so when we bring in a replacement unit, we have to go through the arduous process of resyntonizing, which takes weeks.

DR. HELLWIG:

But would you agree that essentially a reproducible device is adequate if it does over its lifetime basically the same thing (in frequency)?

MR. WARD: Yes. You expressed what I meant, but Harry didn't.

DR. JACQUES VANIER, Laval University:

I would like to make just a few comments here. We have heard about this progress going on in all these fields: quartz, rubidium, hydrogen, cesium. Now this depends somewhat on the scientific interest at the moment, and I was surprised to see today a paper on the rubidium gas cell. And it gave some new evidence that we could control things that people accepted in the past as uncontrollable. They all said, "Well, we have a light shift and that is it. We live with it." And now we see somebody who comes out with the ideas or maybe applies ideas that were old. They were coming from 1960 and we come back to it.

Now the same thing in quartz. These things are all going in parallel and all going further. And it depends quite a lot on the scientific interest of a person like Dr. English, like Dr. Besson, and like Harry Peters for hydrogen masers.

Now let me say something about hydrogen masers. About more than ten years ago, I heard statements like the one you made. The statement: If we had put at the time, ten or twelve years ago, the amount of effort that they have been putting in cesium, you would see where we could have been now. Now where did we go wrong, Harry?

MR. PETERS:

I don't think the ball is in our hands as a matter of fact. As I recall, there were some statements in various meetings about the necessity of making a profit, and some of the research had to be done by selling devices. So whoever got devices at that time got research-built devices.

I don't know of any further enthusiasm in the private area for pursuing this. Now I don't know whether we went wrong. Perhaps some of us should get out of the laboratory; perhaps some of us should stay in it.

DR. HELLWIG:

I am getting signals that the bus is waiting and I apologize for discontinuing the discussion. I would like to thank the panel and the audience for this lively discussion.

TIME DOMAIN MEASUREMENT OF FREQUENCY STABILITY

A tutorial introduction

Jacques Vanier and Michel Têtu

Laboratoire d'Electronique Quantique
Département de Génie Electrique
Université Laval
Québec G1K 7P4
Canada

ABSTRACT

The paper outlines the theoretical basis behind the definition of frequency stability in the time domain. Various types of variances are examined. Their differences and interrelation are pointed out. Systems that are generally used in the measurement of these variances are described.

INTRODUCTION

Radio frequency sources give an output signal which, in general, is affected by small fluctuations in their amplitude, phase or frequency. The nature of these fluctuations may be random or deterministic. Due to the large number of users of these sources and the variety of fields in which they are applied, a problem has arisen in the method of characterization of the frequency fluctuations. In the case of the deterministic fluctuations such as linear frequency drift, it has been found, in most cases, that a specification of the fractional frequency deviation on a per day or per month basis is satisfactory. However, in the case of non deterministic or random fluctuations, a mathematical treatment based on probability concepts is necessary. In the past, depending on the field of interest, several methods of characterization have been used and sometimes have led to confusion.

More recently a proposal has been made by the IEEE Subcommittee on Frequency Stability that the spectral density $S_y(f)$ of the fractional frequency fluctuation, y, and the two sample variance, $\sigma_y(\tau)$, be used to characterize frequency stability respectively in the frequency domain and in the time domain [1]. These two parameters have resulted as a logical conclusion from a rather large amount of theoretical and experimental work on the subject [2],[3]. However, other parameters, specially in the time domain, have been studied and have been found most interesting in specific cases [4],[5],[6],[7]. The subject still raises great interest.

In the present paper we give a brief description of the main theoretical concepts involved in the definition of the frequency stability in the time domain. The interrelation through the spectral density $S_y(f)$ of some of the various variances that have been studied up to now is made explicit. Finally several systems used for the measurement of stability in the time domain are described.

This paper was prepared at the request of the Program Committee of this Meeting. Several excellent reviews and tutorial papers on this subject have been published in recent years [8],[9],[10],[11],[12]. Consequently it appears difficult in the writing of such a tutorial introduction, to avoid repetitions or to improve on all these papers. The most one can do at this stage, is to present the material in a slightly different manner. In particular the present text, specially on the theoretical section, owes a great deal to the recent review by Dr. Rutman [8]; the reader is strongly encouraged to consult that excellent article.

A - THEORY

Definitions

According to a well accepted notation, the instantaneous output voltage of a signal generator can be written as [1]:

$$V(t) = \left[V_0 + \varepsilon(t)\right] \sin\left[2\pi\nu_0 t + \varphi(t)\right] \; , \qquad (1)$$

where V_0 and ν_0 are the nominal amplitude and frequency and $\varepsilon(t)$ and $\varphi(t)$ are amplitude and phase fluctuations. It is assumed that $\varepsilon(t)$ is very small with respect to V_0 and can be entirely neglected. The instantaneous frequency of the oscillator is defined as:

$$\nu(t) = \nu_0 \left(1 + \frac{\dot{\varphi}(t)}{2\pi\nu_0}\right) \; , \qquad (2)$$

where $\dot{\varphi}(t)$ stands for $d\varphi(t)/dt$. We also define the fractional frequency fluctuation as:

$$y(t) \equiv \frac{\dot{\varphi}(t)}{2\pi\nu_0} \; , \qquad (3)$$

and we assume that

$$\left| \frac{\dot{\varphi}(t)}{2\pi\nu_0} \right| \ll 1 \; . \qquad (4)$$

A proposed means to characterize the frequency stability of an oscillator is the spectral density of y denoted by $S_y(f)$. Its dimensions are Hz^{-1}. A measurement giving an estimate of $S_y(f)$ would then characterize the stability of the oscillator in the frequency domain.

This can be done in practice with a spectrum analyzer. However, one must be aware that only estimates of $S_y(f)$ can be obtained because of frequency range limitations and finite observation times.

Experimental studies of various frequency sources have shown that for all practical purposes, the frequency fluctuations spectrum of the most common oscillators can be represented by a truncated polynomial in the Fourier frequency domain:

$$S_y(f) = h^{\alpha} f^{\alpha} , \tag{5}$$

where α is an integer, ranging from -2 to +2. The frequency fluctuations spectral density can be related to the phase fluctuations spectral density through the relation

$$S_y(f) = \frac{f^2}{\nu_0{}^2} S_{\varphi}(f) . \tag{6}$$

The α has been associated with various types of fluctuations either in the phase or frequency representation.

α	Type of fluctuations
2	White Phase
1	Flicker of Phase
0	White Frequency
-1	Flicker of Frequency
-2	Random Walk of Frequency

However, another characterization of the frequency stability can be made by considering that $y(t)$ is a random function of time. Then, a statistical parameter measuring in some sense the excursions of the values y of the random function $y(t)$ around its mean value should characterize the frequency stability of the oscillator. In statistics, the standard deviation σ or the variance σ^2 is often used as statistical parameter. We could define a variance for the instantaneous frequency. In practice the frequency is measured over a time interval ,τ, called the averaging time and the variance is calculated through the relation:

$$\sigma^2(\tau) = \left\langle \left\{ (\overline{y}_i) - \langle \overline{y} \rangle \right\}^2 \right\rangle , \tag{7}$$

where $\langle\ \rangle$ means an average over an infinite number of samples. Without going too deeply into questions of statistical concern, we make a few assumptions. First, we assume that stationnarity applies to our model. By this we mean that a displacement of the time coordinates does not change the statistics of our ensemble. Secondly we assume ergodicity, that is, averages over the ensemble can be replaced by time averages on one of the samples. The sign $\langle\ \rangle$ in equation (7) then becomes a time average. Furthermore, we can assume that $\langle \overline{y} \rangle = 0$ and $\sigma^2(\tau)$

becomes simply $<(\overline{y})^2>$. Since $<(\overline{y})^2>$ implies an infinite time average this variance is an ideal theoretical concept which is commonly called the true variance $I^2(\tau)$. In practice it is clear that one can only do measurements either for a finite time or on a finite number of samples and thus obtain an estimate of this ideal statistical parameter. Furthermore it is found that $I^2(\tau)$ diverges for certain types of noise such as flicker frequency fluctuations. In order to avoid these problems, various scientists have proposed several types of variances obtained from limited amount of samples. As will be explained below, one of them, the "two - adjacent - sample variance" studied by Barnes and Allan, has been proposed as a time domain measurement of frequency stability [1].

An attempt at measuring time domain frequency stability

Before defining the two sample variance let us examine the process of frequency measurement itself. We assume that a digital frequency counter is used to measure the frequency. The measurement is then made over a finite time τ and one obtains an average of the frequency over this time interval τ. In other words, the counter gives the number of cycles n_k during the time interval τ. Figure 1 is an experimental arrangement by which the frequency of oscillator (1) is measured, oscillator (2) being used as the time base of the counter. For the purpose of simplifying the picture let us assume that oscillator (2) is perfect in the sense that its frequency is free of fluctuations. All fluctuations in the measurements would then come from frequency fluctuations of oscillator (1). The counter takes measurements in the sequence shown in figure 2(a). The result for N measurements may be as shown in figure 2(b). Here $\tau = t_{k+1} - t_k$ and

$$\overline{\nu}_{k\tau} = \nu_0 \, (1 + \overline{y}_{k\tau}) \ , \tag{8}$$

$$<\overline{\nu}> \simeq \frac{1}{N} \sum_{i}^{N} \overline{\nu}_{i\tau} \ . \tag{9}$$

The average value of the random variable $\overline{\nu}$ is only an estimate of the actual average frequency, average being done on N samples. One may then calculate for the N samples the variance:

$$\sigma_{\overline{\nu}}^{\ 2} = \frac{1}{N-1} \sum_{j=1}^{N} (\overline{\nu}_j - <\overline{\nu}>)^2 \ . \tag{10}$$

In order to continue the analysis it is assumed that the oscillator does not show systematic drifts with time. If such drifts are present, they are removed from the data and the following analysis applies.

The variance of the random variable $\overline{y}$ may be readily written as:

$$\sigma_y^2(N, T, \tau) = \frac{1}{N-1} \sum_{k=1}^{N} \left(\overline{y}_k - \frac{1}{N} \sum_{j=1}^{N} \overline{y}_j \right)^2 . \tag{11}$$

1. First we note that $\sigma_y^2(N, T, \tau)$ is itself a random variable. It is an estimate of the true variance $I^2(\tau)$ made on N samples. Its average value made on several sets on N samples, $\sigma_y^2(N, T, \tau)$, should be close to the true variance. At the limit where N tends to infinity it should be equal to $I^2(\tau)$.

2. The results of the experiments discussed above can be used to form an histogram as illustrated in figure 3. If the number N of samples is large enough the figure may be rather smooth and a good estimate of $\sigma_y^2(N, T, \tau)$ can be obtained from this curve through a measurement of the half width at half the height. (For a normal distribution $\sigma = 1.17\ (\Delta\overline{y})_{\frac{1}{2}}$).

3. If the experiment is repeated for other averaging times τ, the results obtained with the histogram technique may behave as shown in figure 4. The value of $\sqrt{\sigma_y^2(N, T, \tau)}$ for each of the histograms may then be plotted as a function of τ. The results are shown in figure 5. These results appear very interesting, and give in a sense an indication of the frequency stability of the oscillator in the time domain. However, several difficulties arise when the technique of measurements is changed. For example it is observed that in the region where the variance varies as $1/\tau^2$ an increase in the number of samples N does not alter its value, providing this number of samples is large enough. This is not the case in the region where $\sigma_y^2(N, T, \tau)$ is independent of τ. In that region an increase in the number of samples N shows up by an increase of the variance. One is then faced with a problem of a variance whose values depend on the number of samples. Furthermore, in the region where σ^2 varies as $1/\tau^2$ the value of σ^2 depends on the frequency bandwidth of the measurement system.

The two-sample variance

In order to avoid these problems, and to facilitate intercomparison between the results reported by workers in various fields, one has then to make a choice on N and preferably the ratio T/τ. It has been proposed, following the work of Allan [13]:

1) that the following weighed sample variance be used:

$$\sigma(N, T, \tau) = \frac{1}{N-1} \sum_{i=1}^{N} \left(\overline{y}_i - \frac{1}{N} \sum_{j=1}^{N} (\overline{y}_j) \right)^2 ; \tag{12}$$

2) that $N = 2 \rightarrow$ two-sample variance ;

3) that $\frac{T}{\tau} = 1$, or no dead time between measurements [1]. This variance is a random variable and its average, abreviated

$\sigma_y^2(\tau)$, is given by:

$$\sigma_y^2(\tau) = < \sigma_y^2 (2, T, \tau) > = \frac{1}{2} < (\bar{y}_{k+1} - \bar{y}_k)^2 > . \qquad (13)$$

It is proposed as a characterization of frequency stability in the time domain.

This variance has very interesting properties:

1. It is standardized in the sense that N and T are fixed.

2. It is equal to the true variance $I^2(\tau)$ for white frequency noise. It is equal to $3/2\, I^2(\tau)$ for white phase noise thus close to it.

3. It is convergent for all types of frequency noise normally encountered in oscillators, that is the five power laws mentioned earlier.

4. Although by definition, one is still faced with an average on an infinite number of samples (in this sense $\sigma_y^2(\tau)$ is still an idealization), good estimates of it can be obtained by a relatively limited number of measurements, m , of the pairs:

$$\sigma_y^2(\tau , m) \simeq \frac{1}{m} \sum_{j=1}^{m} \frac{1}{2} (\bar{y}_{k+1} - \bar{y}_k)_j^2 . \qquad (14)$$

For $n > 10$ it has been shown that the confidence interval to be associated to $\sigma_y(\tau)$ in such an estimate if of the order of [14].

Confidence interval = $K_\alpha \, \sigma_y(\tau) / \sqrt{m}$,

where K_α is a constant depending on the power law predominent, but is not far from unity.

5. Finally, tables have been compiled which translates from one type of variance to another in relation to variations in N and the ratio T/τ , and this, for the five power laws commonly encountered in oscillators [15]. Of particular interest is the bias function

$$B_2 (\frac{T}{\tau}, u) = \frac{< \sigma_y^2 (2, T, \tau) >}{< \sigma_y^2 (2, \tau, \tau) >} ,$$

since in general a simple counter will be characterized by a dead time $(T/\tau \neq 1)$ between successive measurements.

The main disadvantages of $\sigma_y^2(\tau)$ are:

1) it diverges for power law spectral densities greater than -2;

2) it does not discriminate between white phase noise and flicker of phase noise.

Relation between the time domain and the frequency domain: other types of variance

The time domain frequency stability, as characterized earlier either through the true variance or the two-sample variance, can be interpreted in a different way. The operation of the counter, averaging the frequency for a time τ may be thought of as a filtering operation. The transfer function , $H(f)$, of this equivalent filter is then the Fourier transform of the impulse response $h(t)$. It can be shown that the time domain frequency stability is then given by [16],[17].

$$<\sigma^2(N, T, \tau)> = \int_0^\infty S_y(f)\,|H(f)|^2\,df\ , \tag{15}$$

where $S_y(f)$ is the spectral density of frequency fluctuations. In the case of the true variance and of the Allan variance, we have:

$$h_I(t) = \begin{cases} 0 & t < -\tau \\ \frac{1}{\tau} & -\tau < t < 0 \\ 0 & t > 0 \end{cases}\ , \qquad h_A(t) = \begin{cases} 0 & t < -\tau \\ \frac{1}{\sqrt{2}\,\tau} & -\tau < t < 0 \\ -\frac{1}{\sqrt{2}\,\tau} & 0 < t < \tau \\ 0 & t > \tau \end{cases}\ , \tag{16}$$

$$H_I(f) = \frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}\ , \qquad H_A(f) = \frac{\sin\frac{2\omega\tau}{2}}{\left(\frac{\omega\tau}{2}\right)^2}\ . \tag{17}$$

These relations are illustrated in figure 6.

This "transfer function approach" has been exploited by several authors to elaborate new types of variances for characterizing oscillator frequency stability in the time domain. It is the equivalent of digital filtering used in data processing.

Hadamard Variance

The sequence of measurements shown in figure 6 for the Allan variance, which consists of two samples ($N = 2$) can be expanded to a sequence of a greater number of samples. A sequence for the case where $N = 10$ is shown in figure 6(e), where the impulse response of the equivalent filter is plotted as a function of t . The variance for this sequence is then [18]:

$$<\sigma_H^2(N, T_D, \tau)> = <(\bar{y}_1 - \bar{y}_2 + \bar{y}_3 - \ldots - y_N)^2>\ , \tag{18}$$

where T_D is the dead time between measurements. The square modulus of the transfer function of this filter is [20]:

$$|H_H(f)|^2 = \left(\frac{\sin \pi \tau f}{\pi \tau f}\right)^2 \left(\frac{\sin N \pi T f}{\cos \pi T f}\right)^2 . \tag{19}$$

It is illustrated in figure 6(f) for no dead time and $N = 10$. The characteristics of this variance, interesting for the topic of frequency stability characterization are as follows:

a) For the case of no dead time ($T_D = 0$), the transfer function of the equivalent filter has a main lobe, centered at $f_1 = 1/2\tau$, and whose width is equal to:

$$f = \frac{\pi^2}{16 \tau N} \text{ (equivalent rectangular filter)}$$

It may thus be made very narrow by increasing N. In this sense this filtering process is well suited for spectral analysis and this property has been exploited to obtain the spectral density of frequency fluctuations using time domain measurements [20]. This is easily seen from relation (15) by realizing that $|H(f)|^2$ can be approximated by a narrow square window over which $S_y(f)$ does not vary much. The spectral density is then given by:

$$S_y(f_1) \simeq \frac{\tau}{N} < \sigma_H^2 (N, \tau, \tau) > \tag{20}$$

This is seen to lend itself to a straightforward computation, in order to obtain an estimate of the spectral density without the use of a spectrum analyser.

b) However, one should be aware of severe limitations in this technique. There exist secondary side lobes in $H_H(f)$, which appear at harmonics of $f_1 = 1/2\tau$, for no dead time between measurements. These can be minimized by proper adjustment of the dead time between measurements or by proper weighing of the samples of the measurements. The properties of the transfer function have been well studied in the case of a weighing by the binomial coefficients and by a pseudo-sinusoïdal function [18], [19], [20].

Modified sample variance

Boileau and Picinbono have introduced a variance which can be interpreted in the case of its digital realization as a modified sample variance. [6] Their relations, when translated into the notation generally adopted in the field of frequency stability gives a variance as follows:

$$\sigma^2_{mod} (N, T, \tau) = \bar{y}_{(N+1)/2} - \frac{1}{N} \sum_{i=1}^{N} \bar{y}_i)^2 , \tag{21}$$

where N is odd. Thus $y_{(N+1)/2}$ is the central sample of the set of N

samples. Rutman has examined the case when $N=3$ and $T=\tau$ [8]. The measurement sequence and the square modulus of the transfer function are shown in figure 6(g) and 6(h) respectively. The variance is given by:

$$\langle \sigma^2_{mod}(3,\tau,\tau)\rangle = \int_0^\infty S_y(f)\,\frac{16}{9}\,\frac{\sin^6 \pi\tau f}{(\pi\tau f)^2}\,df \quad . \tag{22}$$

The modified sample variance has the advantage of being convergent for all five power law spectral densities examined up to now, plus two others, where $\alpha=-3$ and $\alpha=-4$. However, it does not discriminate between white phase noise and flicker phase noise better than $\sigma_y(\tau)$.

A special case of the sample variance

It is general practice to study the frequency stability or measure the variance as a function of averaging time τ. This is what is done in the sample variance described above. In opposition to this practice, De Prins and Cornelissen [7] have studied frequency fluctuations over intervals T for fixed averaging times τ. One can then in principle study the variance $\sigma_y{}^2(T)$ of the instantaneous frequency $(\tau\to 0)$ as a function of the time interval T. This variance is very different from the one described previously. In the present case all values of $y(t)$ are scanned as T is varied in opposition to the averaging over τ considered up to now.

In the case where τ does not tend to zero but is fixed at a given value, the variance then becomes a special case of the sample variance and the behaviour of $\sigma_y{}^2(T)$ can be obtained from the bias function, B_2.

The high pass variance

A close look at equation (15) suggests that $\sigma^2(\tau)$ can actually be defined through the transfer function $H(f)$ of the equivalent filter corresponding to the measurement sequence. Rutman has suggested that this approach could be taken even if the actual measurement sequence was not existing [5]. Then, $H(f)$ could be given the shape desired. Of course the inverse Fourier transform of $H(f)$ is not necessarily a step wise function that could be implemented in a straightforward manner by a counting technique. Other measurement techniques have then to be implemented.

In this approach the variance is written:

$$\sigma^2(\tau) = \frac{1}{\pi^2\nu_0{}^2\tau^2}\int_0^\infty S_\varphi(f)\,|H_\varphi(f)|^2\,df \quad , \tag{23}$$

where $S_\varphi(f)$ is the phase spectral density and is related to $S_y(f)$ through relation (6). The variance is then defined in terms of $H_\varphi(f)$ and not in terms of the measurement sequence.

From equation (23) one sees that for the Allan variance, the square modulus of the phase transfer function of the equivalent filter is

$$|H_{\varphi A}(f)|^2 = \sin^4 \pi f \tau \text{ . (See figure (7))} \tag{24}$$

This is essentially a high pass filter, having an oscillating nature with a period $1/\tau$. Low frequency components $f < (\pi\tau)^{-1}$ are filtered out, this being an essential character of the Allan variance. Consequently it appears that $|H_\varphi(f)|^2$ could be essentially a high pass filter and essentially the same character for the variance would be obtained. In fact, calculations show that, when a second order high pass filter, with cut off frequency $f_c = (\pi\tau)^{-1}$, is used to calculate a so called "high pass variance", the general behaviour with the power law spectral density is essentially the same as the behaviour of the Allan variance. Both variances have the same asymptotic slopes with τ and both variances cannot differentiate between white phase noise and flicker phase noise.

Band pass variance

Following this line of thought and recognizing the nature of the limitations of the high pass variance, Rutman [8] has suggested that a bandpass filter be used for $|H_\varphi(f)|^2$ with a center frequency equal to $(1/2\tau)$ and a constant Q factor, say equal to 1. In that case the behaviour of $\sigma_{BP}(\tau)$ is quite different from that of $\sigma_y(\tau)$ or $\sigma_{HP}(\tau)$; it shows complete discrimination between the five power law models in its asymptotic behaviour as a function of τ.

Of course, the method for measuring $\sigma_{BP}(\tau)$ is not a conventional one incorporating a frequency counter. One uses a phase comparator (loose-phase-lock technique), a bandpass filter and a r.m.s. voltmeter. In this sense, it is the same type of system as the one used in reference [17] and essentially falls in the class of systems used for studying frequency stability in the frequency domain. It appears natural to think of frequency stability measurements in the time domain as being done through a time sequential technique and a statistical analysis of the resulting data. This should be kept in mind in the practical implementation of systems designed for the measurement of frequency stability in the time domain.

An unified approach

In the previous sections, various types of variance were examined. The approach taken has been one in which the measurement sequence was identified; the transfer function of the equivalent filter implementing the impulse response for the sequence in question was established, and the variances could be calculated through relation (15). This method is very useful in pointing out the limit of utilization of a particular variance in respect to the power law frequency model and also in understanding the reason of these limits.

These variances however have all been introduced as particular cases for special needs. Recently, Lindsey, Chie, Leavitt and Lewis [21], [22],[23] have introduced in the picture a method of analysis called the "structure function approach", which emphasizes the fundamental ties between these variances rather than their differences.

The k^{th} average frequency fluctuation over time τ can be written as a difference of phase:

$$\bar{y}_k = \frac{1}{2\pi \nu_0 \tau} \left[\varphi(t_k + \tau) - \varphi(t_k) \right] \quad . \tag{25}$$

We may define the first difference or first increment in phase as:

$$\Delta^{(1)} \varphi(t_k, \tau) = \varphi(t_k + \tau) - \varphi(t_k) \quad . \tag{26}$$

The frequency difference $(\bar{y}_{k+1} - \bar{y}_k)$ appears as a second difference in phase:

$$(\bar{y}_{k+1} - \bar{y}_k) = \frac{1}{2\pi \nu_0 \tau} \left[\varphi(t_k + \tau) - 2\varphi(t_k) + \varphi(t_k - \tau) \right] \quad , \tag{27}$$

and the second phase increment is defined as:

$$\Delta^{(2)} \varphi(t_k, \tau) = \varphi(t_k + \tau) - 2\varphi(t_k) + \varphi(t_k - \tau) \quad . \tag{28}$$

It is readily realized that the first difference in phase is used in the definition of the true variance, while the second difference in phase is used in the definition of the Allan variance (two samples). Lesage and Audoin [20] have proposed to continue the process further and have obtained an expression for the n^{th} difference in phase, which includes the binomial coefficient as a weighing factor. This analysis has led them to the implementation of the Hadamard variance in which the measurement sequence is weighed by the binomial coefficients.

From this it appears that a common basis may be expected under the definition of the various variances examined above. The rank of the phase increment appears to play a major role. In the approach of Lindsey and Chie this point is stressed. The N^{th} phase increment is defined as:

$$\Delta^{N} \varphi(f, \tau) = \sum_{k=0}^{N} (-1)^k \binom{N}{k} \varphi(t + (N-k)\tau) \quad , \tag{29}$$

where $\binom{N}{k} = \frac{N!}{k!\,(N-k)!}$ (Binomial Coefficient) . (30)

The structure function of phase is then defined as:

$$D^{(N)}(\tau) = < (\Delta^{(N)} \varphi(t, \tau) > \quad . \tag{31}$$

Stationnarity of the N^{th} difference is assumed in the wide sense. It is then shown that the variances defined earlier can all be expressed in terms of these structure functions:

True variance

$$I^2(\tau) = \frac{1}{(2\pi\,\nu_0\,\tau)^2}\, D_{\varphi}^{(1)}(\tau) \quad ; \tag{32}$$

Allan variance

$$\sigma_y^{\,2}(\tau) = \frac{1}{2(2\pi\,\nu_0\,\tau)^2} \quad D_{\varphi}^{(2)}(\tau) \quad ; \tag{33}$$

Modified three-sample variance

$$< \sigma^{\,2}_{y\,\mathrm{mod}}(\,3,\,\tau,\,\tau\,) > \; = \frac{1}{9(2\pi\,\nu_0\,\tau)^2}\, D_{\varphi}^{(3)}(\tau) \quad ; \tag{34}$$

Hadamard variance (weighed by the binomial coefficients)

$$< \sigma^{2}_{\mathrm{H.B.C.}}(\,N,\,\tau,\,\tau\,) > \; = \frac{1}{(2\pi\,\nu_0\,\tau)^2}\, D_{\varphi}^{(N)}(\tau) \quad ; \tag{35}$$

This approach thus clearly shows that the variances utilized up to now by various authors have a common basis, in occurence, a structure function of phase. On the other hand, this structure function is related to the spectral density through the relation:

$$D_{\varphi}^{N}(\tau) = 2^{2(N-1)}(2\pi\,\nu_0\,\tau)^2 \int_0^{\infty} S_y(f)\, \frac{\sin^{2N} \pi\,\tau\,f}{(\pi\,\tau\,f)^2}\, df \quad , \tag{36}$$

which effectively, as stressed earlier, provides means for evaluating the spectral density $S_y(f)$ as filtered with a transfer function

$$H_{sf}(f) = \frac{\sin^{2N} \pi\,\tau\,f}{(\pi\,\tau\,f)^2}\; x,\; 2^{\,2(N-1)} \tag{37}$$

through sequential measurements in the time domain.

In the previous paragraphs a structure function of phase was introduced as a means for describing frequency stability in the time domain. A structure function of frequency, however, can also be introduced to describe frequency stability; it is written $D_y^{(N)}(\tau)$. Lindsey and Chie [22] have shown, by studying the mathematical differences between these structure functions, that the true variance is essentially a measure of phase instability while the Allan variance is a measure of frequency instability. This type of reasoning has led them to suggest that a function of the product of the two types of variance could be a parameter by which frequency stability in the time domain could be characterized.

It should be pointed out that the characterization of the stability of an oscillator could in principle be made through tables of the structure functions $D^{(N)}(\tau)$. The user could calculate from these tables the type of variance that is best suited for his particular application. In a sense, these structure functions can be thought of, as characterizing completely the frequency stability of the oscillator in the time domain in the same sense as $S_y(f)$, does it in the frequency domain.

Long time frequency fluctuations

In real oscillators it is possible to observe very long term frequency fluctuations, that is, very slow fluctuations which may appear over periods of days, months and years. These may originate either from slow random fluctuations or from deterministic drifts in the behaviour of the oscillator.

Slow random fluctuations

The above analysis was limited to five power laws of the spectral density $(-2 < \alpha < +2)$. Very slow frequency fluctuations predominate at very low frequencies and are thus represented by more negative slope power laws such as $\alpha = -3$ or -4. The greatest negative slope that the Allan variance can handle is $\alpha = -2$, a random walk of frequency type of noise; for more negative slopes it diverges. Since direct spectral analysis of these slow fluctuations is not experimentally feasible, it appears that the other types of variance mentioned earlier may be useful. In fact the modified three sample variance $\sigma^2_{y\,mod}(3, \tau, \tau)$ converges for f^{-3} and f^{-4} types of noise with respective slope τ^2 and τ^3. Consequently for very slow frequency fluctuations one may then have to use a variance different from the two sample variance in order to have meaningfull interpretation of time domain data.

Deterministic drifts

Systematic drifts are generally observed in oscillators. These drifts may be represented by a polynomial [9]. A model for fractional frequency drifts is:

$$y(t) = d_1 t^1 + d_2 t^2 + \ldots d_n t^n \quad . \tag{38}$$

For the first term, representing a linear frequency drift, the Allan variance varies as τ^2. It is time dependent for higher order drifts. The other variances mentioned above consequently are found very useful in characterizing these higher order polynomial drifts. In particular the modified three sample variance varies as τ^2 for quadratic frequency drifts while the Hadamard variance weighed by binomial coefficients varies as τ^3 for cubic frequency drifts. The behaviour of the asymptotic value of $\sigma(\tau)$ versus τ for various power laws of spectral density and various orders of frequency drifts are summarized in table 1. It should be under-

stood that the structure function approach can also be applied to analyze these long term frequency fluctuations, although emphasis has not been placed on this point in this paper.

In this section we have examined the properties of several types of variances that can be used for characterizing frequency stability of oscillators in the time domain. Differences between these variances have been shown; their interrelation through the fractional frequency fluctuations spectral density has also been emphasized. In fact, it has been shown that these variances, in the case of random fluctuations, are essentially elegant means of representing the spectral density $S_y(f)$ through parameters which can be measured with simple systems implemented with a frequency counter and a calculator for doing statistical analysis. Such systems will now be described.

B - TIME DOMAIN MEASUREMENT SYSTEMS

Ideally, frequency stability measurements require a frequency reference much more stable than the oscillator to be studied. A system realizing this condition can be implemented easily for the measurement of the frequency stability of most common oscillators. However, for stable, state-of-the art, oscillators, it is necessary that the reference oscillator be at least as stable as the oscillator studied. Frequency stability measurements in the time domain can be done by two different methods which are related to the detection of two different parameters: frequency and phase.

One can determine the mean frequency over finite observation times and calculate, for a given number of samples, a certain variance as described in the previous section. Although these measurements involve the well developed technology of frequency or period counting, their use is limited by the fact that they give information only on the mean frequency and the frequency fluctuations. Furthermore, in many systems, the measurement samples are not adjacent in time which, in some cases, alter the value of the statistical parameter calculated.

Many experimental set up's have been considered during the last several years, as the technology evolved (see for example reference [24], [25] and [26]. We shall limit our discussion to a few of them illustrating their main principles.

The most simple set up is the "Direct Frequency Counting System" shown in figure 8. In this case the reference oscillator is the frequency counter time base. It is suitable for the study of low performance oscillators. The output of the counter gives the mean frequency over a preselected time interval. Its recording on a trip chart, magnetic tape or in a digital memory allows one to use best estimate curve fitting methods, in order to find any systematic frequency trends or drifts. These drifts are removed prior to statistical analysis. The two sample variance (Allan

variance) or any other desired variances are calculated from the set of corrected data. The main limitation of this set up is the ± one count and the accuracy of the time base reference.

The "Frequency Heterodyne Technique", shown in figure 9, is another important set up used when the oscillator to be studied is more stable than the frequency counter time base. One must use another oscillator as the reference. First, the oscillator frequency is translated down to a value which can be conveniently measured by the counter. The translation is realized by mixing the oscillator signal with the reference signal set at a convenient different frequency and by detecting only the difference frequency (beat frequency). A synthesizer can be very helpful in many cases. The optional link between the reference and the counter allows an increase in the counter performance and a precise measurement of the absolute frequency. Statistical analysis is done on the beat frequency as described for the former set up. One must refer the fluctuations to the nominal frequency of the oscillator.

If the oscillator and the reference frequencies are very close, one may multiply each one by a different factor, creating a more sizeable beat frequency. Such a system is shown in figure 10. Frequency counting and statistical analysis are achieved as in the previous set up. Unfortunately, noise may be introduced in the multiplier chains which set limits to this technique.

Measurements of stability in the frequency domain require the detection of phase or frequency fluctuations. While equipped to do so, the same set up can be used to do measurements in the time domain. These experimental systems are called "Phase Locked Reference Systems". A voltage controllable reference oscillator is phase-locked to the oscillator signal. When loosely locked, as indicated in figure 11, the phase detector delivers a signal which is proportional to the phase difference between the two oscillators. If the reference is considered more stable than the oscillator, the fluctuations of this signal is attributed to the random changes of the oscillator phase. When this signal is processed through a differentiator, a low frequency signal, proportional to the frequency fluctuations, is obtained. It is then possible to use a voltage to frequency converter driven by this signal, to generate a low frequency oscillation, fluctuating in the same manner as the frequency of the original oscillator. Average frequencies, over time intervals τ, are determined by a counter and statistical analysis is performed as in the previous systems.

If the reference oscillator is tightly locked in phase to the oscillator studied, as shown in figure 12, the command signal, applied to the reference is proportional to the frequency changes between the two oscillators. Again if the reference oscillator is much more stable than the other oscillator, this command signal fluctuates in the same way as does the frequency of the oscillator. Statistical analysis of

this signal is performed with the set up described in the loose phase locked reference case.

The two systems act simply as frequency translators with non unity conversion factor. The sensitivity of these systems is enhanced by a factor proportional to the ratio of the nominal oscillator frequency to the nominal frequency of the converter. Time domain regions, where the results are significant, are determined by the servo loop characteristics. Usually a second order loop is used in order to achieve optimum performance; the parameters to adjust are the natural frequency and the damping factor [27],[28],[29]. The loose Phase Locked Reference System is normally used when frequency stability for averaging times below 1 sec is needed, while the tight Phase Locked Reference System is preferred for longer averaging times.

Obviously these systems are well suited for the measurement of phase or frequency fluctuations in the frequency domain. For this type of measurement, it is necessary to measure the spectral density at the phase detector in the first system or the spectral density of the command signal in the second system. Since for highly stable oscillators the information lies in a spectrum containing frequencies much lower than one hertz, a very low frequency spectrum analyser is required. Digital real time spectrum analysers are indicated for such measurement but they are expensive. The method of bandpass filtering proposed by Rutman can be applied to give measurements either in the time domain or in the frequency domain with these systems [5].

All the experimental systems described previously deal with the measurement of frequencies or periods averaged over a finite observation time, τ. Characterization of frequency stability is done through statistical analysis on an ensemble of these averaged frequencies. It was pointed out in the theoretical section that the frequency stability of an oscillator can also be characterized by a measurement of phase differences. To do so, one has to measure the phase of an oscillator and to calculate phase differences, spaced in time by an interval, τ. Measurement techniques of phase are well developed in the field of time scale implementation since a time scale can be graduated in terms of phase with 2π radians as a unit of time (one period). Thus, measurements of phase differences correspond to time difference measurements.

We will describe two systems for measuring frequency stability by time differences. They are the Dual Mixer Time Difference System [30] and the Phase Modulated Phasemeter [31]. Both systems give about the performance of ±1 picosecond time interval resolution and a limit of resolution of roughly $1 \times 10^{-13}\,\tau^{-1}$ when used to measure fractional frequency stability. Since these systems measure the phase of oscillators, the evaluation of the two-sample variance can be done without dead time and no correction factors are needed. Furthermore the systems can be used to compare oscillators of exactly the same nominal frequency (clock oscil-

lators).

The Dual Mixer Time Difference System is illustrated in figure 13. The two oscillators are of the same type and have the same nominal frequency, ν_o. A common oscillator translates down to ν_b each oscillator frequency in two identical channels. Phase comparison is done between the two beat signals. Zero crossing detection on each signal is achieved and serves to trigger pulse generators. The time interval between each pair of pulses is a measure of the relative phase of the two oscillators; the system acts as a sophisticated phase detector. Any fluctuations of this time interval can be considered as fluctuations of the phase of one or both oscillators. A time interval counter measures these phase (time) differences and statistical analysis is done according to prescribed theoretical calculations.

The translation has the effect of increasing the system resolution by a factor which is approximately ν_0/ν_b. When used to characterize the frequency stability in the time domain, the sampling time is the period or a multiple integer of the beat signal. Consequently a synthesizer used as common oscillator becomes a very convenient tool. Noise contribution from the common oscillator is greatly reduced when the nominal phase shift between the two beat signals is small. An adjustable phase shifter is then placed in series with one oscillator in order to satisfy this condition. (See the Appendix for a calculation on that system.)

The Phase Modulated Phasemeter System is illustrated in figure 14. Again two identical oscillators are compared in phase. Each signal frequency is multiplied, then mixed and filtered to get a beat signal. A non zero frequency beat is generated by modulating the phase of one oscillator with a low frequency signal. Zero crossing detection of this signal gives pulses which can be located in time when compared to a reference signal triggered by the modulating signal. This time interval is a measurement of the phase of one oscillator compared to the other. If one oscillator is considered as a reference, the time interval will be a measurement of the phase of the other oscillator. As in the case of the Dual Mixer Time Difference System, the frequency stability is calculated from this time interval recording for different observation times by simple phase differences and statistical weighing. With this system, the measurement resolution is increased by the up conversion factor.

The various measurement systems described show a certain hierarchy in the parameters evaluated. When one has access to a signal proportional to the phase, he is allowed to calculate any combination of phase differences, then, any variances. In fact, it is possible to calculate all the structure functions of the phase fluctuations which gives a complete measurement of the frequency stability in the time domain. This is possible

with the two last systems and the loose Phase Locked Reference System when the phase is recorded instead of being differentiated. All the other systems described give access only to frequency and its fluctuations, and this limits the amount of statistical information which can be obtained.

APPENDIX: EXPRESSION OF THE FREQUENCY STABILITY IN THE TIME DOMAIN FOR THE DUAL MIXER TIME DIFFERENCE SYSTEM

The phases at the output of oscillator 1 and oscillator 2 are :

$$\Phi_1(t) = \omega_1 t + \theta_1 + \varphi_1(t) \quad ,$$

and $$\Phi_2(t) = \omega_2 t + \theta_2 + \varphi_2(t) \quad ,$$

where ω_1 and ω_2 are the angular frequencies, θ_1 and θ_2 are the initial phase offsets and $\varphi_1(t)$ and $\varphi_2(t)$ are the phase fluctuations. The phase at the output of the common oscillator is :

$$\Phi_R(t) = \omega_R t + \theta_R + \varphi_R(t)$$

with the corresponding meaning for each parameter or variable.

At the input of each mixer, the phase of the oscillator signals can be represented by the equations already given where θ includes the phase shift added by the variable shifter and $\varphi_1(t)$ and $\varphi_2(t)$ include any phase fluctuations added by the transmission links. The phase of each reference signal can be written as :

$$\Phi_{R,1}(t) = \omega_R t + \theta_{R,1} + \varphi_{R,1}(t)$$

$$\Phi_{R,2}(t) = \omega_R t + \theta_{R,2} + \varphi_{R,2}(t)$$

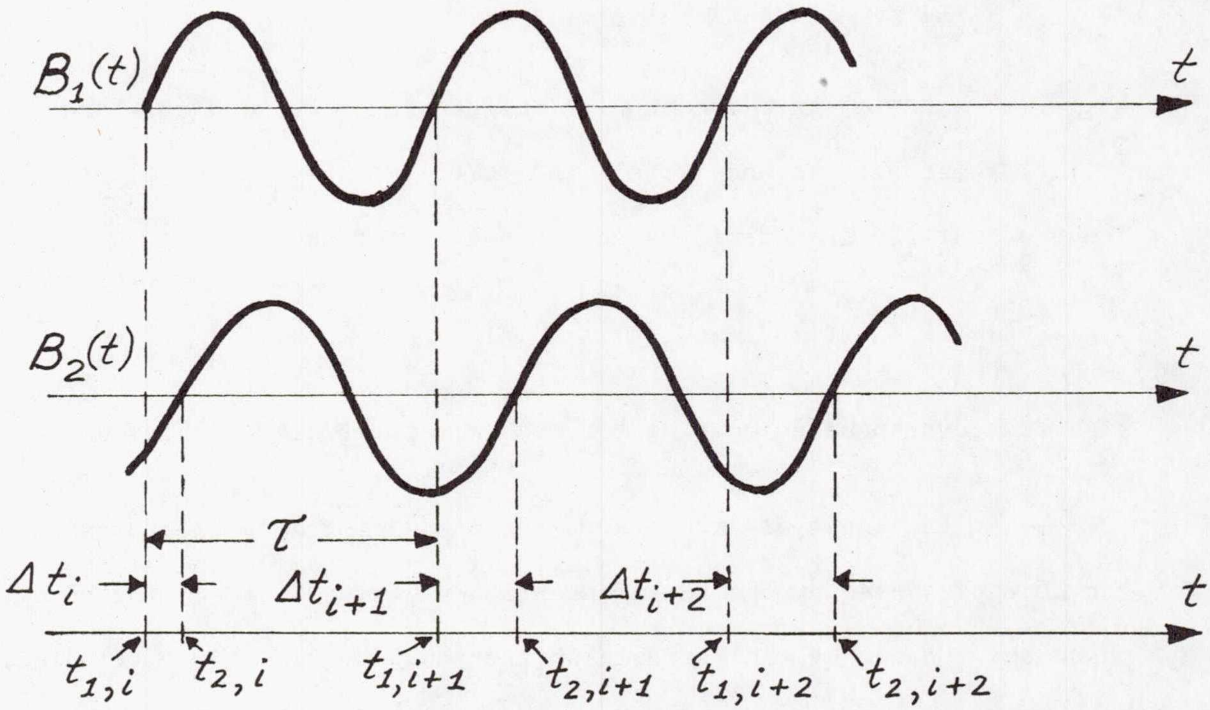
in order to account for any phase shifts and phase fluctuations introduced by the isolation amplifiers and the transmission links. We will see below how these phase perturbations can be made negligeable in the measurement system.

The output of each mixer delivers low frequency beats whose phase are :

$$\Phi_{B,1}(t) = (\omega_R - \omega_1)t + (\theta_{R,1} - \theta_1) + \varphi_{R,1}(t) - \varphi_1(t)$$

$$\Phi_{B,2}(t) = (\omega_R - \omega_2)t + (\theta_{R,2} - \theta_2) + \varphi_{R,2}(t) - \varphi_2(t)$$

These two signals have a time evolution represented as follow :



The positive zero crossings give time events that are a measurement of the relative phase between the two oscillators.

At a certain time, $t_{1,i}$, the phase of the first beat signal is such that the signal is zero; then

$$(\omega_R - \omega_1)t_{1,i} + \theta_{R,1} - \theta_1 + \varphi_{R,1}(t_{1,i}) - \varphi_1(t_{1,i}) = m\,2\,\pi$$

At a certain time $t_{2,i}$, the phase of the second beat signal corresponds to the same criterion :

$$(\omega_R - \omega_2)t_{2,i} + \theta_{R,2} - \theta_2 + \varphi_{R,2}(t_{2,i}) - \varphi_2(t_{2,i}) = (m+n)2\pi$$

where m and n are integers.

Now if we impose that: $\omega_2 \equiv \omega_1$ and define $\omega_B = \omega_R - \omega_2 = \omega_R - \omega_1$, we obtain the identities

$$\omega_B t_{2,i} + \theta_{R,2} - \theta_2 + \varphi_{R,2}(t_{2,i}) - \varphi_2(t_{2,i}) = (m+n)\,2\pi$$

$$\omega_B t_{1,i} + \theta_{R,1} - \theta_1 + \varphi_{R,1}(t_{1,i}) - \varphi_1(t_{1,i}) = m\,2\pi$$

and by subtraction :

$$\omega_B(t_{2,i} - t_{1,i}) + \theta_{R,2} - \theta_{R,1} - \theta_2 + \theta_1 + \varphi_{R,2}(t_{2,i}) - \varphi_{R,1}(t_{1,i})$$
$$- \varphi_2(t_{2,i}) - \varphi_1(t_{1,i}) = n\,2\pi$$

In this expression, $\theta_{R,2}$ and $\theta_{R,1}$ are constants different by a value introduced by the isolation amplifiers and cable lengths. It is possible to define a constant phase offset,

$$\Delta\theta = \theta_{R,2} - \theta_{R,1} - \theta_2 + \theta_1$$

which can be adjusted by the phase shifter. By doing so, we also set the nominal value of $t_{2,i} - t_{1,i}$; for small time offsets, the two phase fluctuation terms coming from the reference oscillator are correlated and their difference is negligeable when compared to the phase fluctuation difference of the two oscillators. Within this approximation, the time difference becomes :

$$\Delta t_i = t_{2,i} - t_{1,i} = \frac{\varphi_2(t_{2,i}) - \varphi_1(t_{1,i})}{\omega_B} - \frac{\Delta\theta}{\omega_B} + \frac{n\,2\pi}{\omega_B}$$

If we look at the next pair of zero crossings we obtain, in a similar way, the time difference :

$$\Delta t_{i+1} = t_{2,i+1} - t_{1,i+1} = \frac{\varphi_2(t_{2,i+1}) - \varphi_1(t_{1,i})}{\omega_B} - \frac{\Delta\theta}{\omega_B} + \frac{n2\pi}{\omega_B}$$

and for the following pair

$$\Delta t_{i+2} = t_{2,i+2} - t_{1,i+2} = \frac{\varphi_2(t_{2,i+2}) - \varphi_1(t_{1,i+2})}{\omega_B} - \frac{\Delta\theta}{\omega_B} + \frac{n2\pi}{\omega_B}$$

The definition of the two sample variance in term of the second difference of phase is :

$$\sigma^2(\tau) = \frac{1}{2}\left(\frac{\varphi(t_{i+2}) - 2\varphi(t_{i+1}) + \varphi(t_i)}{2\pi\nu_0\tau}\right)^2$$

where the averaging time, τ , is the time interval between t_i and t_{i+1} or the time interval between two successive phase measurements of each oscillator. Such a linear combination of phase fluctuations can be obtained by grouping the time differences just derived. Then

$$\Delta t_{i+2} - 2\Delta t_{i+1} + \Delta t_i = \frac{\varphi_2(t_{2,i+2}) - 2\varphi_2(t_{2,i+1}) + \varphi_2(t_{2,i})}{\omega_B} - \frac{\varphi_1(t_{1,i+2}) - 2\varphi_1(t_{1,i+1}) + \varphi_1(t_{1,i})}{\omega_B}$$

If the two oscillators are statistically independent, we can write :

$$(\Delta t_{i+2} - 2\Delta t_{i+1} + \Delta t_i)^2 = \left(\frac{\varphi_2(t_{2,i+2}) - 2\varphi_2(t_{2,i+1}) + \varphi_2(t_{2,i})}{\omega_B}\right)^2 + \left(\frac{\varphi_1(t_{1,i+2}) - 2\varphi_1(t_{1,i+1}) + \varphi_1(t_{1,i})}{\omega_B}\right)^2$$

and in terms of the two sample variance of each oscillator :

$$(\Delta t_{i+2} - 2\Delta t_{i+1} + \Delta t_i) = 2 \left(\frac{\nu_0}{\nu_B}\right)^2 \tau^2 \left[\sigma_2^2(\tau) + \sigma_1^2(\tau)\right]$$

This relation is the corner stone of the Dual Mixer Time Difference System. It shows that a linear combination of the time interval measurements give a value proportional to the two sample variance of each oscillator. When the two oscillators are identical, $\sigma_2^2(\tau) = \sigma_1^2(\tau)$, and the calculated value is twice the value for each oscillator. If one oscillator is much more stable than the other oscillator, $\sigma_1^2(\tau) << \sigma_2^2(\tau)$, the calculated value is directly the value for the instable oscillator. In these calculations, we consider three successive pairs of time event; they are then spaced in time by an interval: $\tau = (\nu_B)^{-1}$, the beat period. It is also possible to skip a determined number of zero crossings between each pair of time events. In this case, the observation time is a multiple integer of $(\nu_B)^{-1}$.

A similar type of calculations applies to the phase modulated phasemeter.

REFERENCES

1. J.A. Barnes, A.R. Chi, S.L. Cutler, D.J. Healey, D.B. Leeson, T.E. McGunigal, J.A. Mullen, W.L. Smith, R.L. Sydnor, R. Vessot and G.M.R. Winkler, "Characterization of Frequency Stability", IEEE Trans. Instrum. Meas. IM 20, pp. 105-120, 1971.

2. IEEE-NASA Symposium on Short Term Frequency Stability, Goddard Space Flight Center, Greenbelt MD, 1964 (Document NASA SP-80).

3. Special Issue on Frequency Stability, Proc. IEEE 54, 1966.

4. R.A. Baugh, "Frequency Modulation Analysis with the Hadamard Variance", Proc. 25th Annual Frequency Control Symposium, Atlantic City, N.J., pp. 222-225, 1971.

5. J. Rutman, "Characterization of Frequency Stability: A Transfer Function Approach and its Application to Measurement via Filtering of Phase Noise", IEEE Trans. Instrum. Meas. IM 23, pp. 40-48, 1974.

6. E. Boileau and B. Picinbono, "Statistical Study of Phase Fluctuations and Oscillator Stability", IEEE Trans. Instrum. Meas. IM 25, pp. 66-75, 1976.

7. J. De Prins and G. Cornelissen, "Power Spectrum, Frequency Stability and Flicker Noise", Proc. Frequency Standards and Metrology Seminar, Laval University, Quebec, Canada, pp. 368-387, 1971 (unpublished).

8. J. Rutman, "Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress", Proc. IEEE 66, pp. 1048-1075, 1978.

9. G.M.R. Winkler, "A Brief Review of Frequency Stability Measures", Proc. 8th Annual PTTI Applications and Planning Meeting U.S. Naval Research Laboratory, Washington, D.C., pp. 489-527, 1976.

10. J.A. Barnes, "A Review of Methods of Analyzing Frequency Stability", Proc. 9th Annual PTTI Applications and Planning Meeting (NASA Goddard Space Flight Center Greenbelt MD), 1977.

11. D.A. Howe, "Frequency Domain Stability Measurements: A tutorial Introduction", NBS Technical Note 579, 1976.

12. G. Rovera, "Stabilita di Frequenza", Lezioni tenuto presso la Scuola di Perfezionamento in Fisica dell'Universita di Pisa, 1974 (Pisa Editrice Tecnica Scientifica).

13. D.W. Allan, "Statistics of Atomic Frequency Standards", Proc. IEEE 54, pp. 221-230, 1966.

14. P. Lesage and C. Audoin, "Characterization of Frequency Stability: Uncertainty due to the Finite Number of Measurements", IEEE Trans. Instrum. Meas. IM 22, pp. 157-161, 1973 (Corrections on this paper have been published later in the issues March 1974 and September 1976 of the same Journal).

15. J.A. Barnes, "Tables of Bias Functions B_1 and B_2 for Variances Based on Finite Samples of Processes with Power Law Spectral Densities", NBS Technical Note 375, 1969.

16. L.S. Cutler and C.L. Searle, "Some Aspects of the Theory and Measurement of Frequency Fluctuations in Frequency Standards", Proc. IEEE 54, pp. 136-154, 1966.

17. R. Vessot, L. Mueller and J. Vanier, "The Specification of Oscillator Characteristics from Measurements Made in the Frequency Domain", Proc. IEEE 54, pp 199-207, 1966.

18. R.A. Baugh, "Frequency Modulation Analysis with the Hadamard Variance", Proc. 25th Annual Frequency Control Symp., Atlantic City, N.J., pp. 222-225, 1971.

19. G. Sauvage et J. Rutman, "Analyse Spectrale du Bruit de Fréquence des Oscillateurs par la Variance Hadamard", Ann. Télécom. 28, pp. 301-314, 1973.

20. P. Lesage et C. Audoin, "Instabilité de Fréquence des Oscillateurs - Mesure dans le Domaine Fréquence par une Méthode Temporelle", L'Onde Electrique 55, pp. 82-89, 1975.

21. W.C. Lindsey and J.L. Lewis, "Modeling, Characterization and Measurement of Oscillator Frequency Instability", Report CDRL item No B020, Lincoln Corporation Pasadena, Cal. 91104, 1974.

22. W.C. Lindsey, C.M. Chie and W.E. Leavitt, "Interpretation and Application of Oscillator Instability Measures", Proc. 8th Annual Precise Time and Time Interval Applications and Planning Meeting, U.S. Naval Research Laboratory, Washington, D.C., pp. 535-560, 1976.

23. W.C. Lindsey and C.M. Chie, "Theory of Oscillator Instability Based Upon Structure Functions", Proc. IEEE 64, pp. 1652-1666, 1976.

24. L.S. Cutler, C.L. Searle, "Some Aspects of the Theory and Measurement of Frequency Fluctuations in Frequency Standards", Proc. IEEE, Vol 54, No 2, p. fig. , February 1966.

25. "Time and Frequency: Theory and Fundamentals", B.E. Blair, editor, NBS Monograph 140, U.S. Dept. of Commerce, Washington, D.C., May 1974.

26. "Frequency and Time", P. Kartaschoff, Academic Press, Monograph in Physical Measurement, London, England, 1978.

27. F.L. Walls, S.R. Stein, "Design Considerations in State-of-the-Art Signal Processing and Phase Noise Measurement Systems", Proc. 30th Annual Symposium on Frequency Control, Electronic Industries Ass., pp. 269-274, June 1976.

28. J. Vanier, M. Têtu, L.G. Bernier, "Transfer of Frequency Stability from an Atomic Frequency Reference to a Quartz Crystal Oscillator", Proc. 32nd Annual Symposium on Frequency Control, Electronic Industries Ass., pp. 520-526, June 1978.

29. "Phase-Locked Loops: Application to Coherent Receiver Design", A. Blanchard, Wiley-Interscience publication, 1976.

30. D.W. Allan, H. Daams, "Picosecond Time Difference Measurement System", Proc. 29th Annual Symposium on Frequency Control, Electronic Industries Ass., pp. 404-411, May 1975.

31. A. Gabry, B. Dubouis, D. Landréa, G. Faucheron, "Dispositif de mesure de la stabilité des horloges étalons, utilisant un phasemètre à modulation de phase", Note Technique NT/TCR/DEF/1606, Centre National des Etudes en Télécommunications, 92131 Issy-les-Moulineaux, Mai 1978.

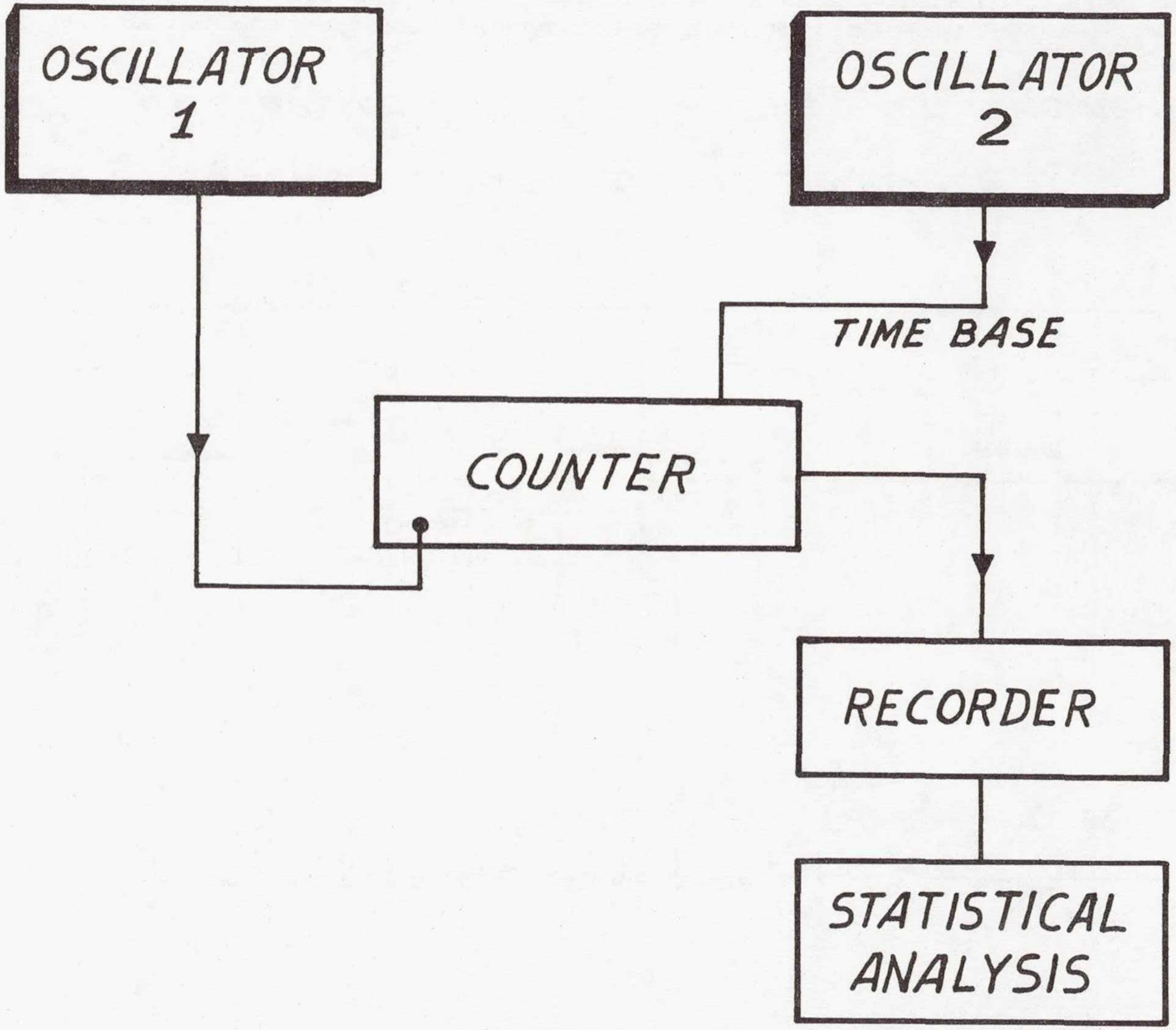


Figure 1. Block diagram of the ideal experimental set up used to measure the frequency of oscillator (1) over an averaging time τ. The frequency of oscillator (2) is assumed to be free of fluctuations.

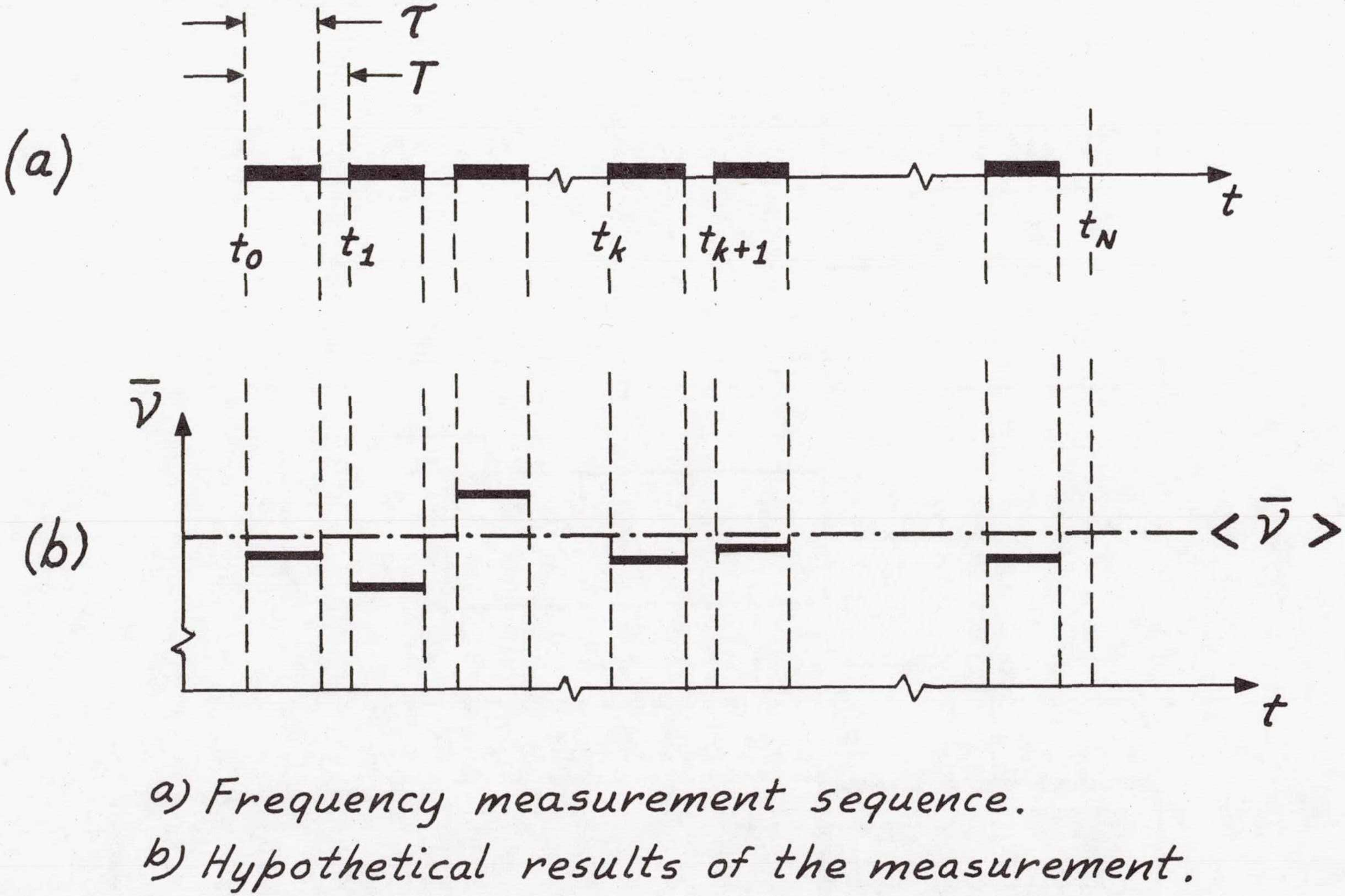


Figure 2. (a) Time sequence used in the measurement of the frequency of an oscillator.
(b) Hypothetical result of the measurements of the frequency of oscillator (1) in the set up of figure (1).

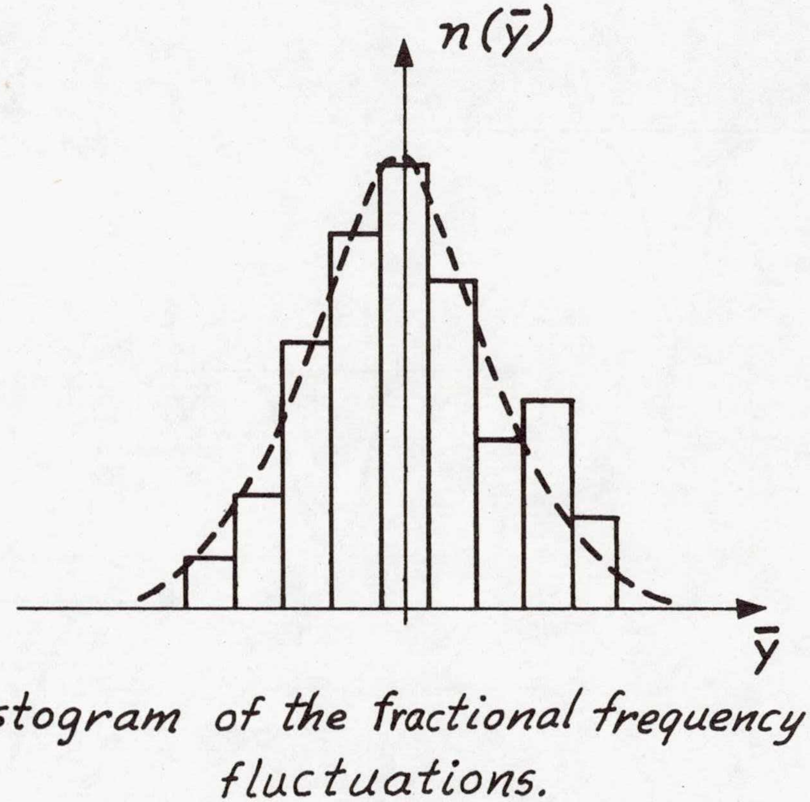


Figure 3. Histogram of the hypothetical results of the frequency fluctuations shown in figure 2.

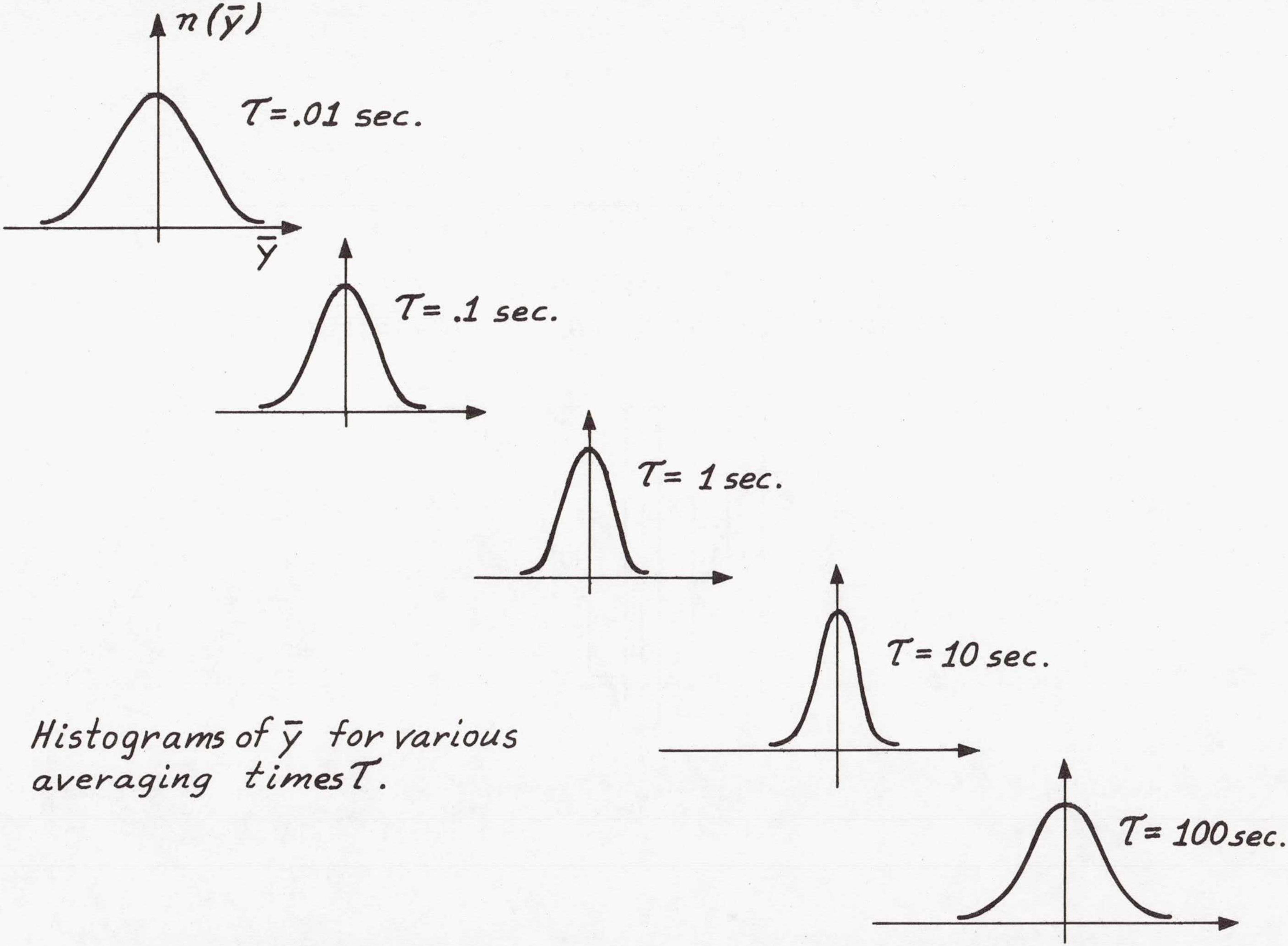


Figure 4. Hypothetical histograms obtained by repeating the measurements of frequency on N samples for various averaging time τ.

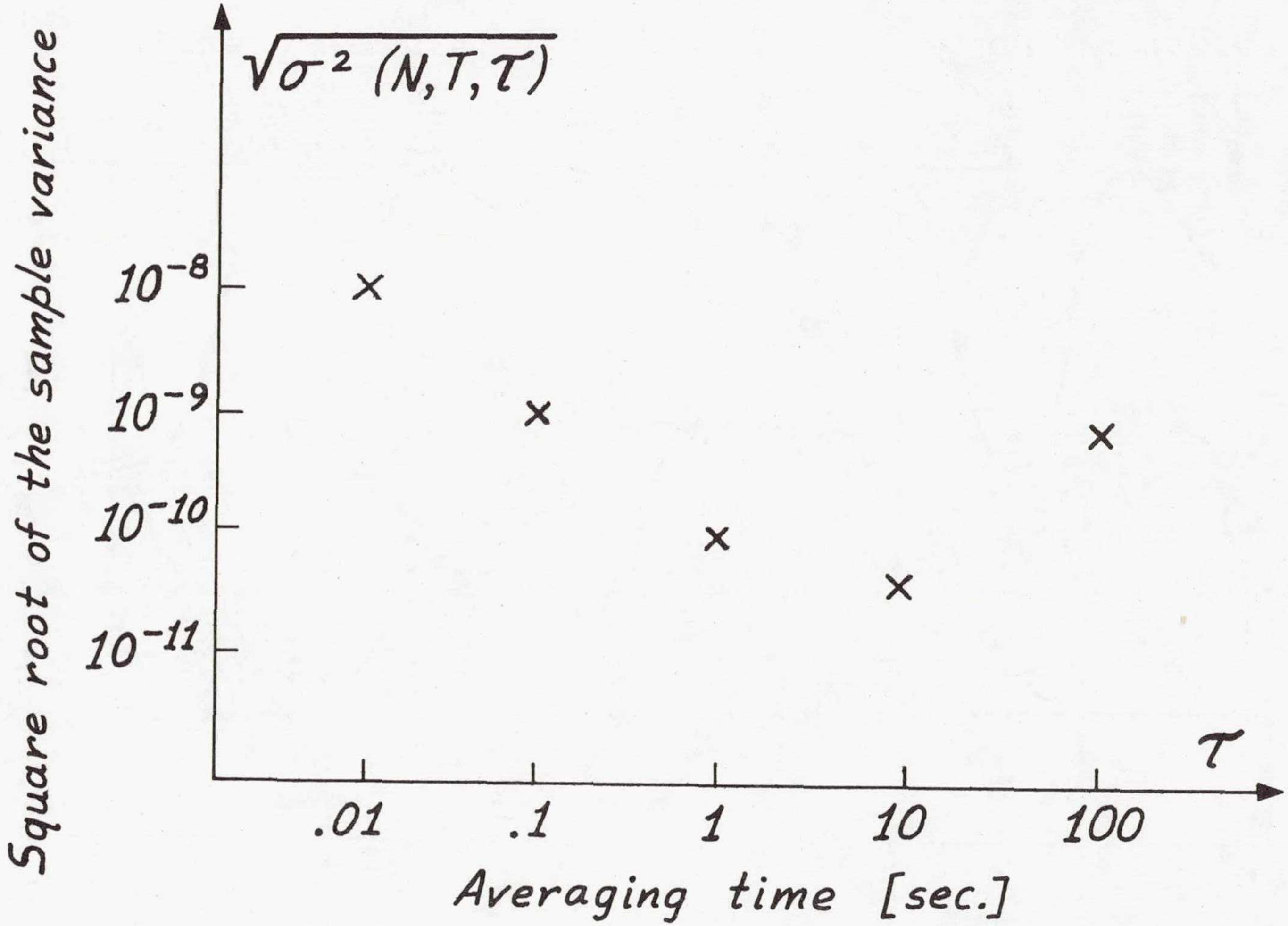


Figure 5. Plot of the square root of the sample variance $\sigma^2(N, T, \tau)$ (standard deviation) obtained for the hypothetical results of figure 4, as a function of the averaging time.

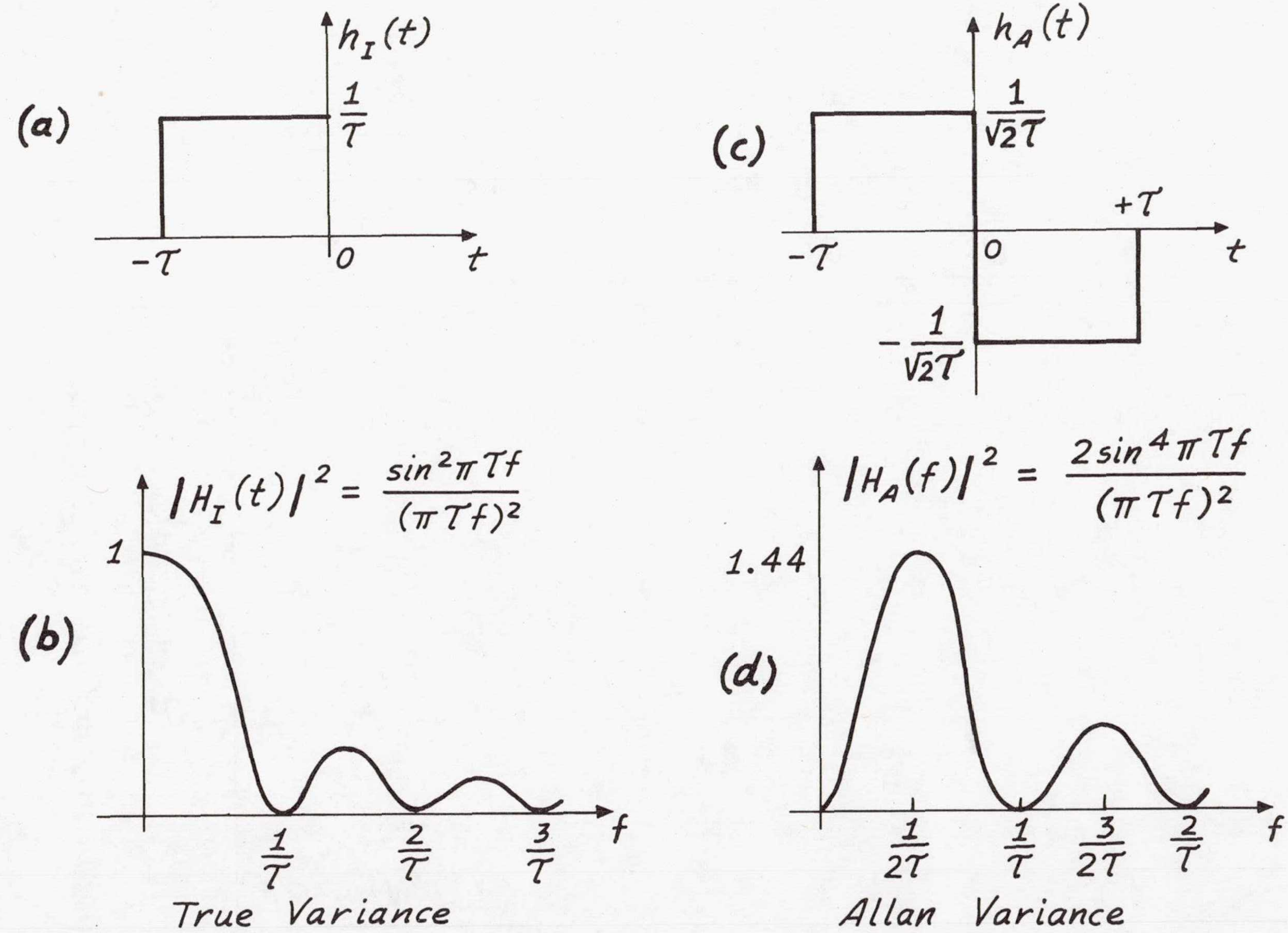


Figure 6. Impulse response h(t) and square modulus of the transfer function H(f) of the equivalent filters used in the calculations of variances (a) and (b): True variance; (c) and (d): Allan variance; (e) and (f): Hadamard variance; (g) and (h): Modified sample variance.

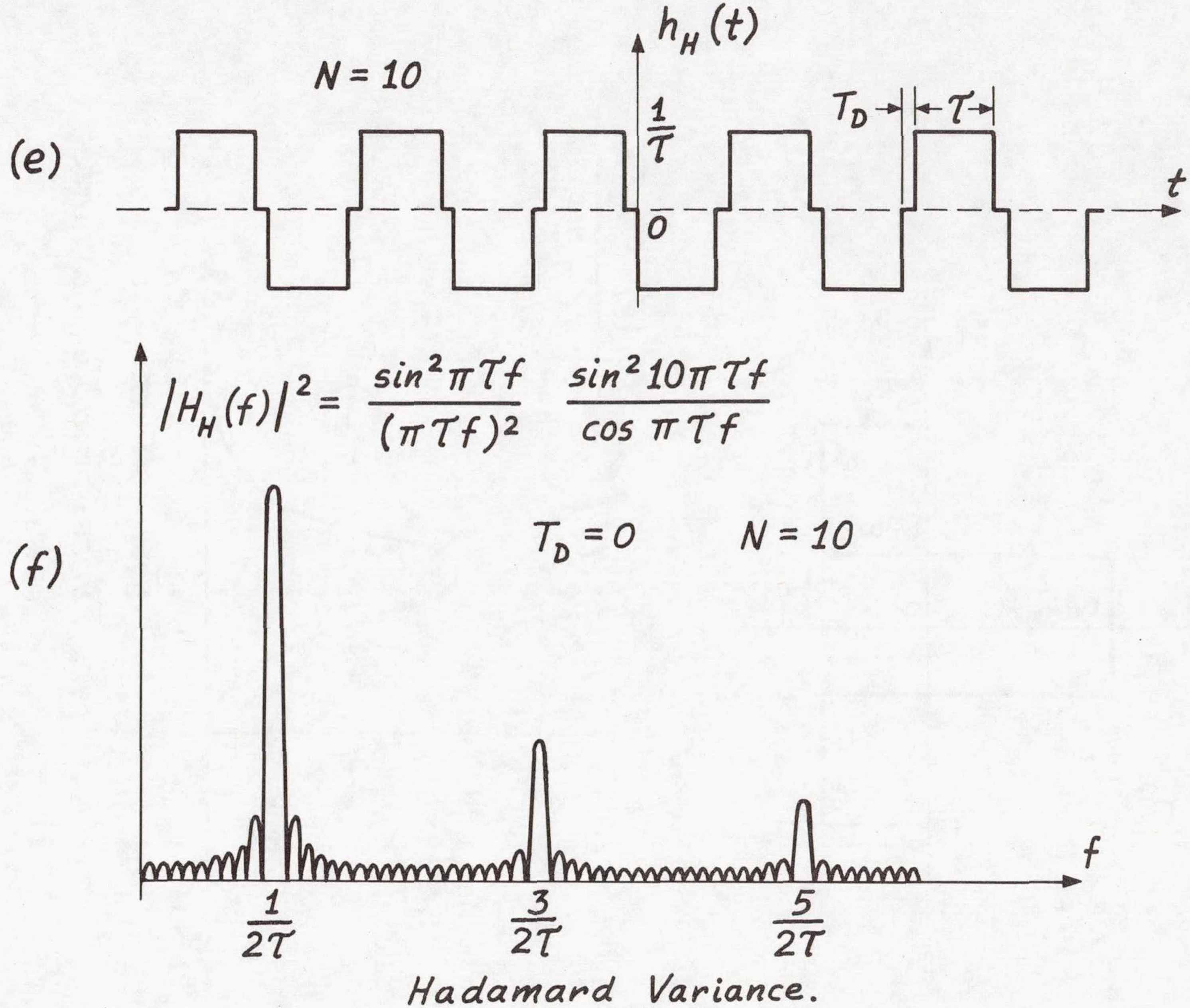
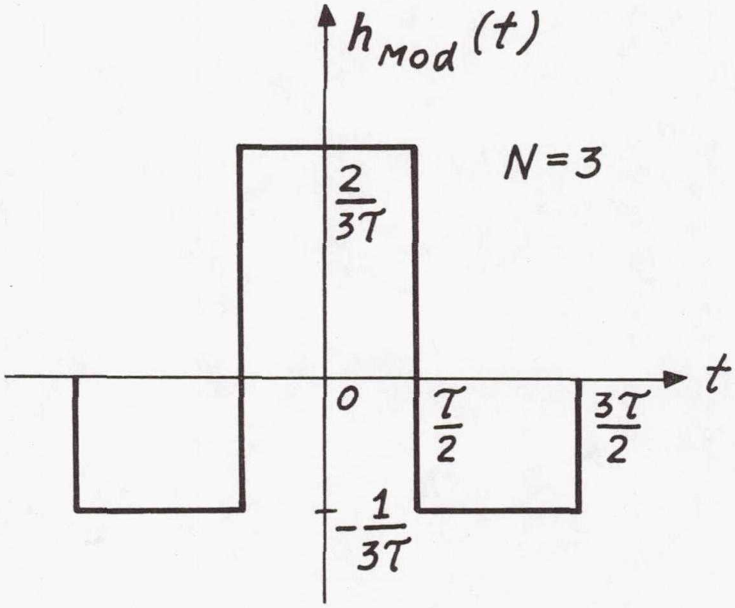
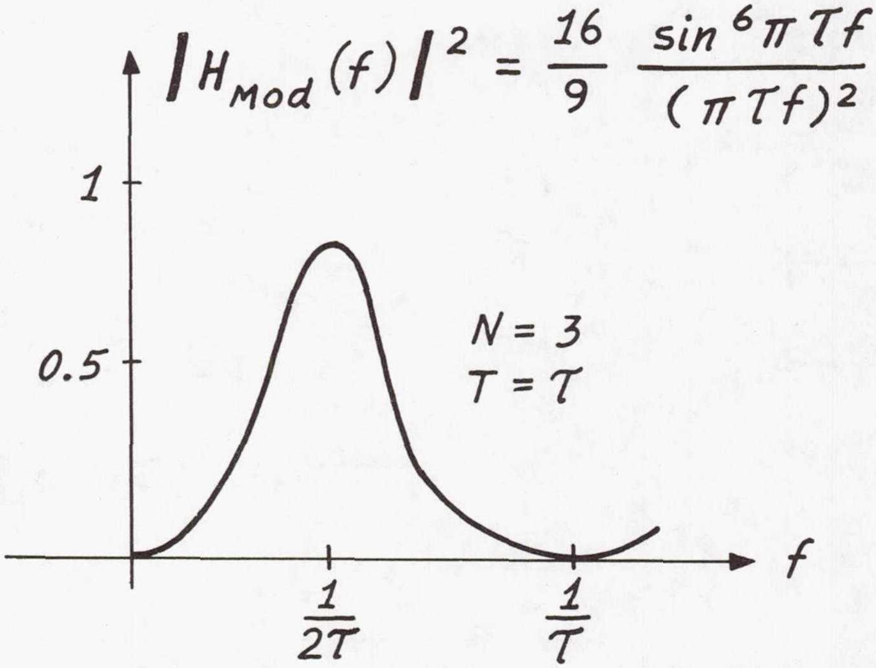


Figure 6 (Cont.)

(g)

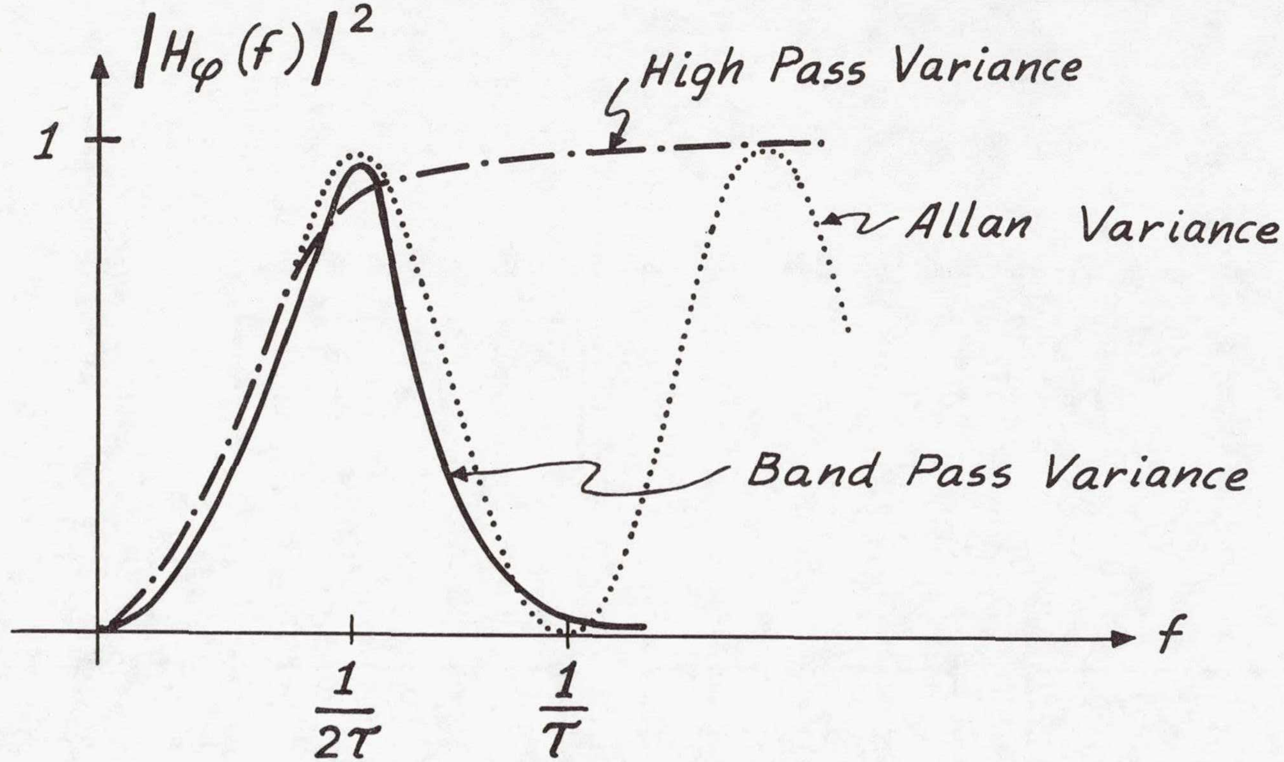


(h)



Modified Three-Sample Variance.

Figure 6 (Cont.)



$$\sigma^2(\tau) = \int_0^\infty S_y(f)\,|H(f)|^2\,df$$

$$\sigma^2(\tau) = \frac{2}{(\pi \nu_0 \tau)^2} \int_0^\infty S_\varphi(f)\,|H_\varphi(f)|^2\,df$$

High Pass and Band Pass Variances.

Figure 7. Square modulus of the phase transfer function $H_\varphi(f)$, for the high pass, low pass and Allan variances.

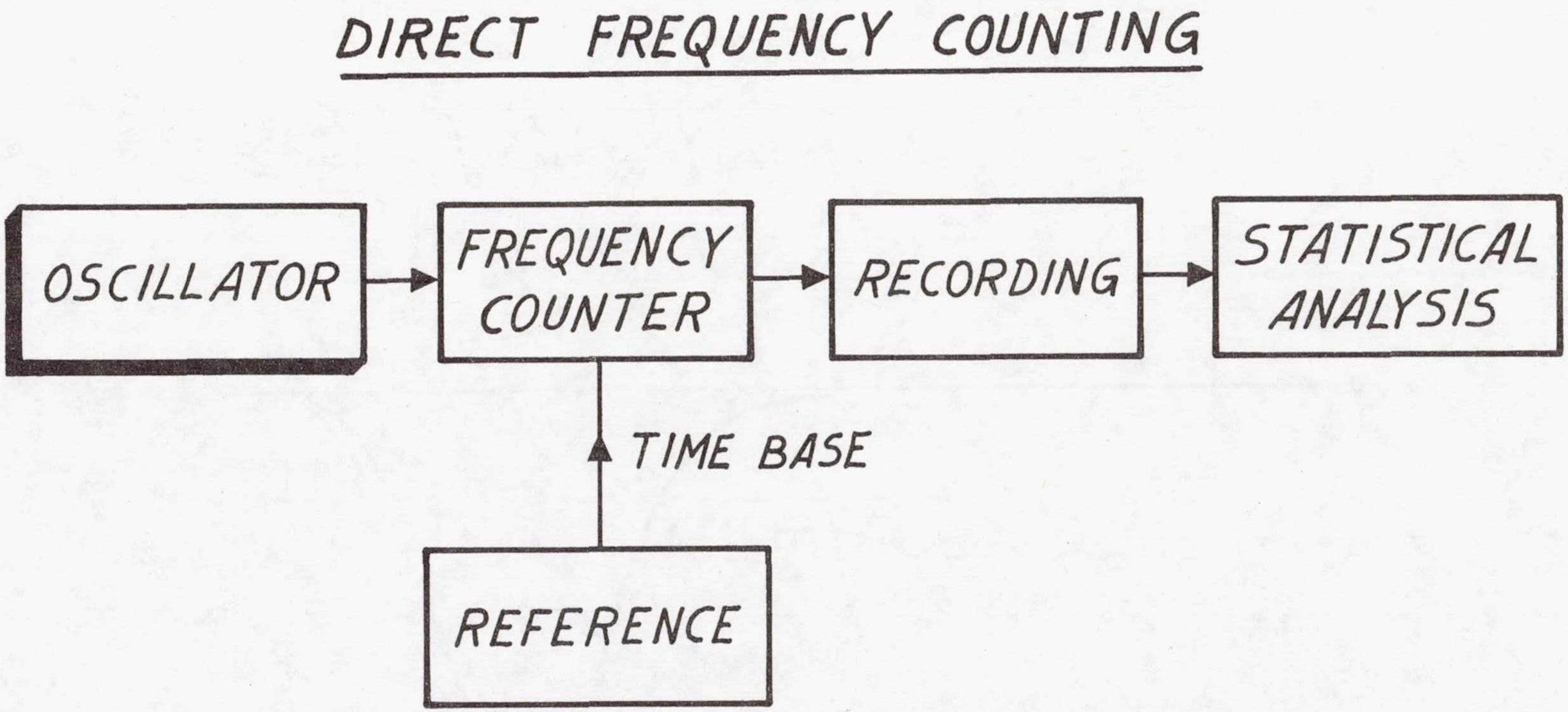


Figure 8. Block diagram illustrating the Direct Frequency Counting method.

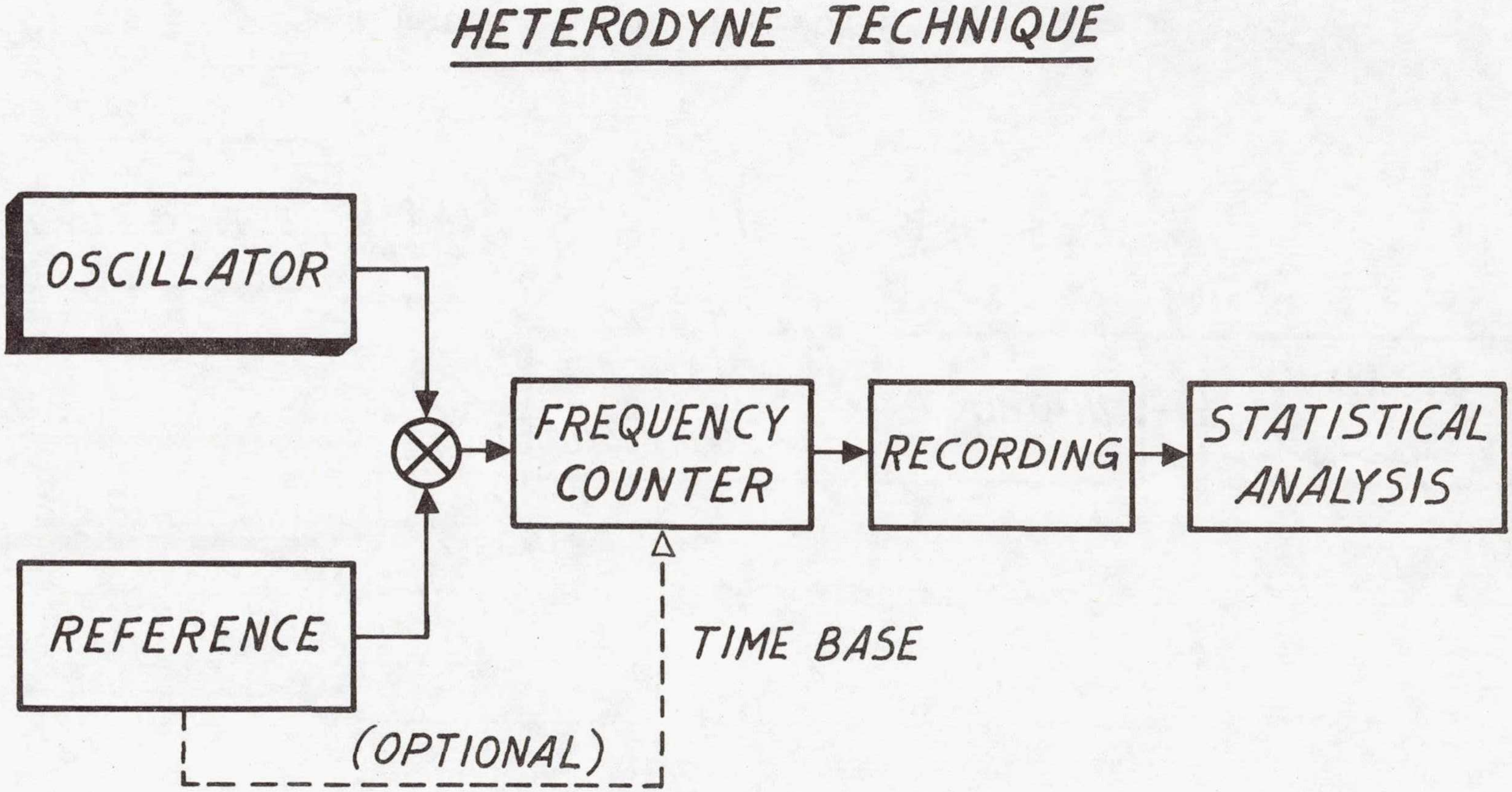


Figure 9. Block diagram illustrating the Heterodyne Technique.

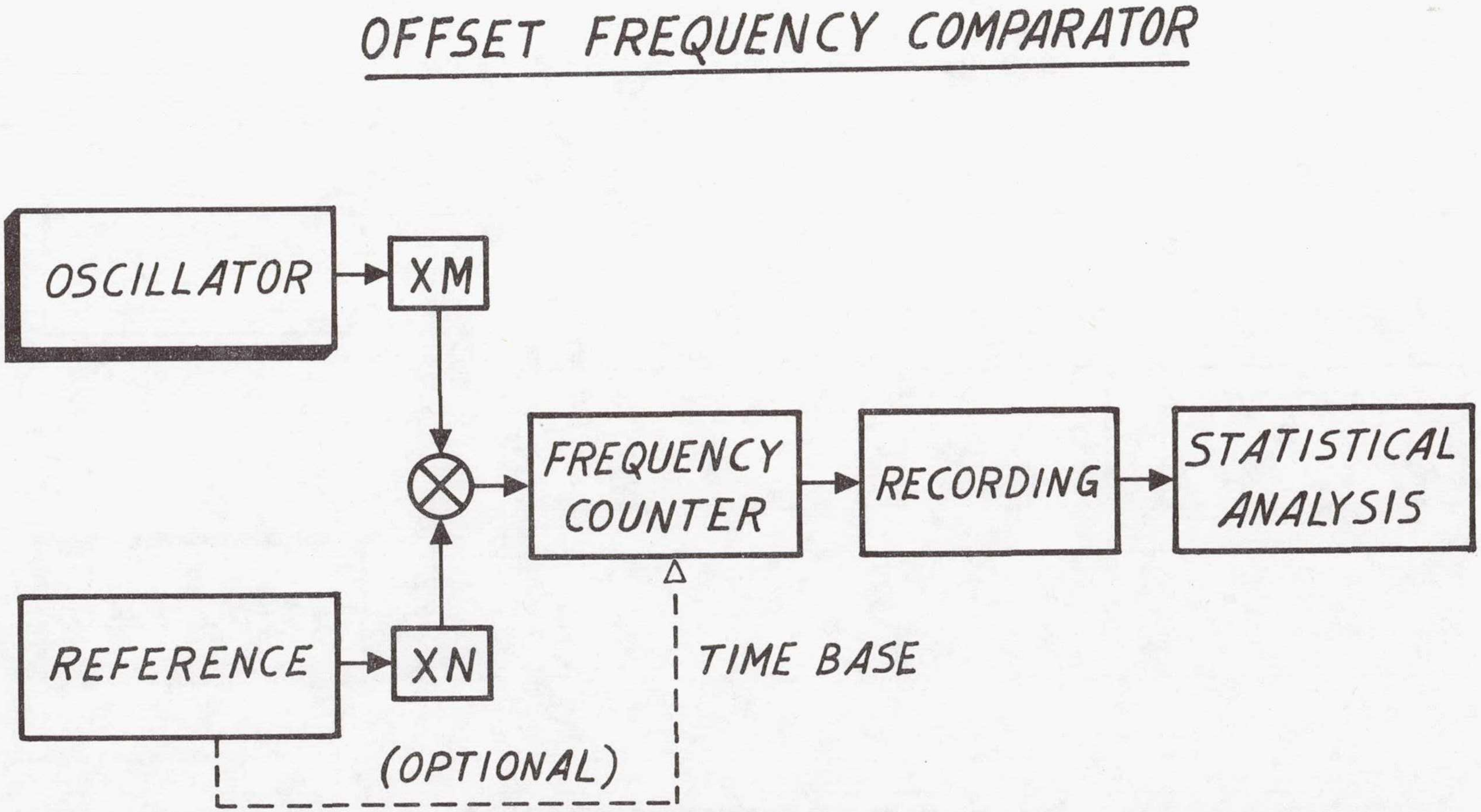


Figure 10. Block diagram illustrating the Offset Frequency Comparator method.

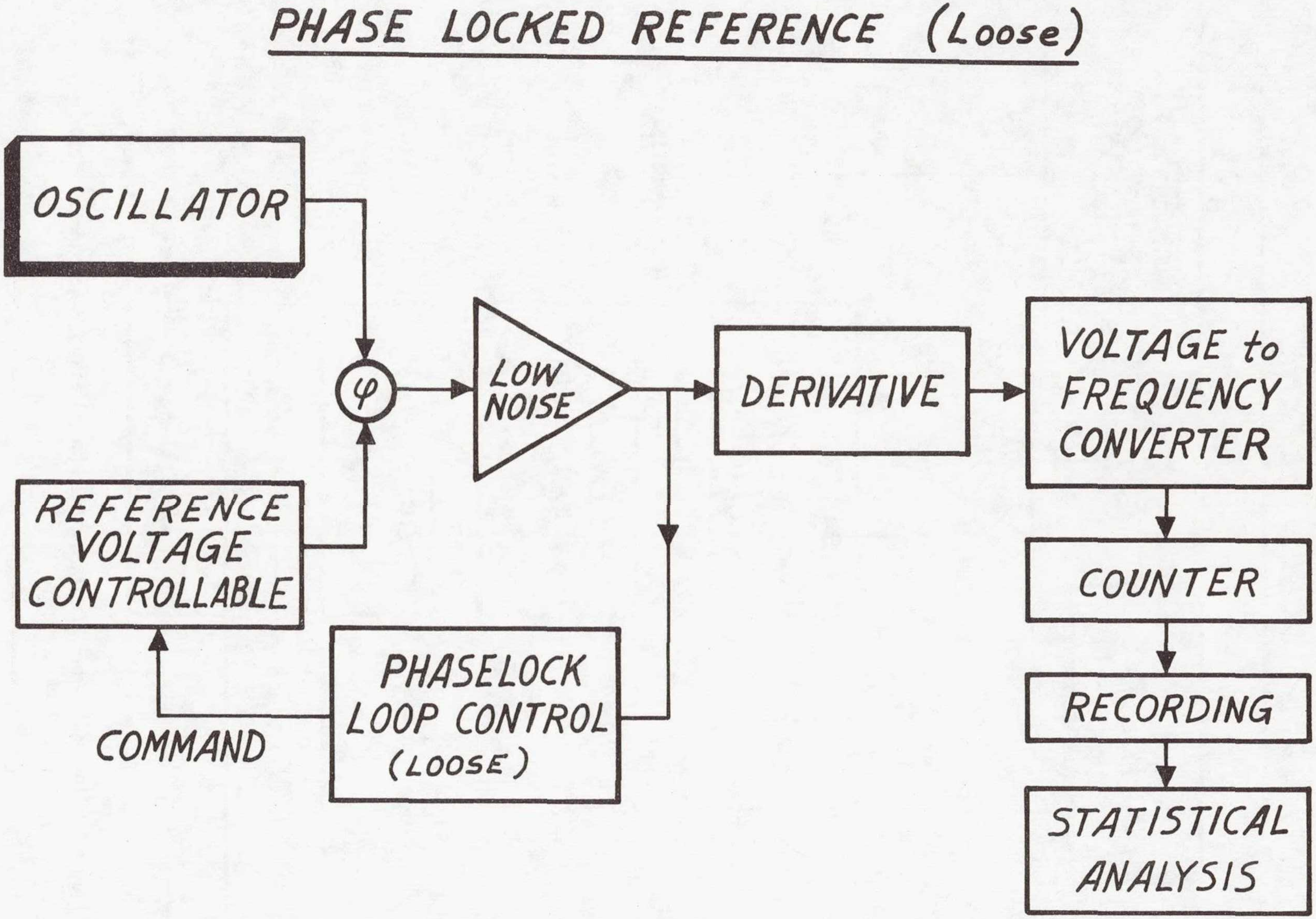


Figure 11. Block diagram illustrating the Phase Locked Reference System. In this system the reference oscillator is loosely locked.

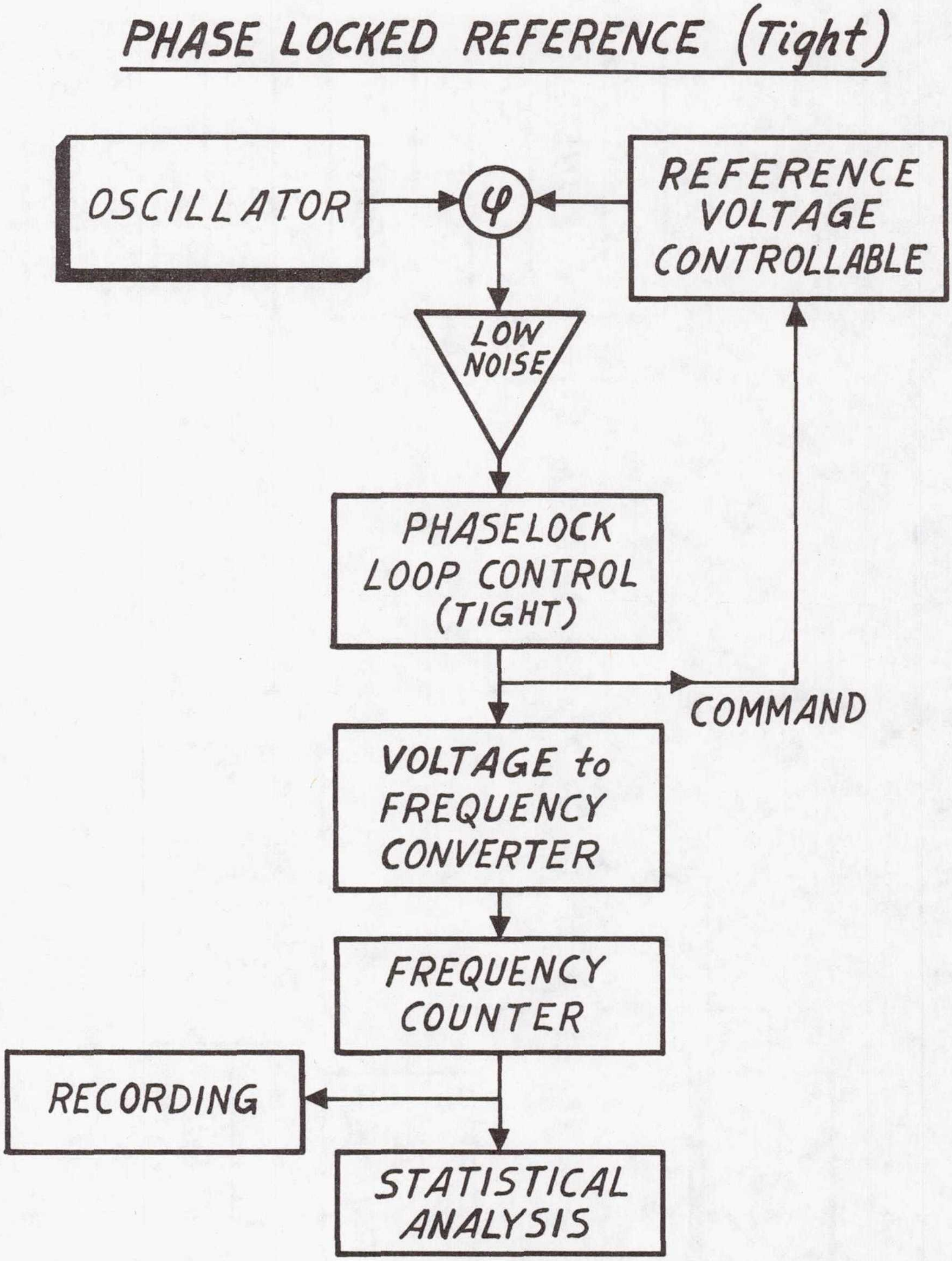


Figure 12. Block diagram illustrating the Phase Locked Reference oscillator in the tightly locked mode.

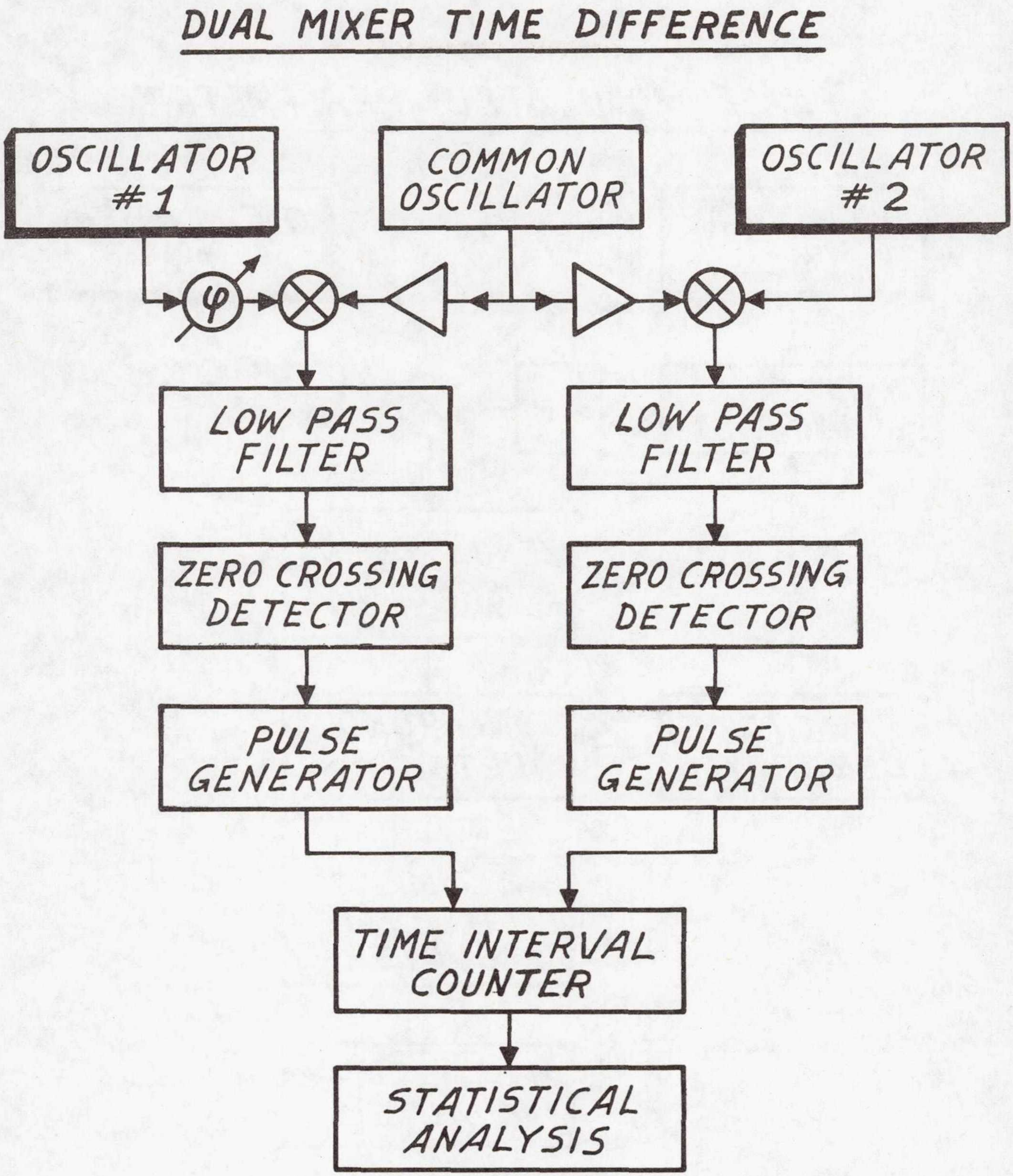


Figure 13. Block diagram illustrating the Dual Mixer Time Difference System. When the two oscillators have tendency to move off frequency they must be phase locked loosely together. This loop is not shown on this figure.

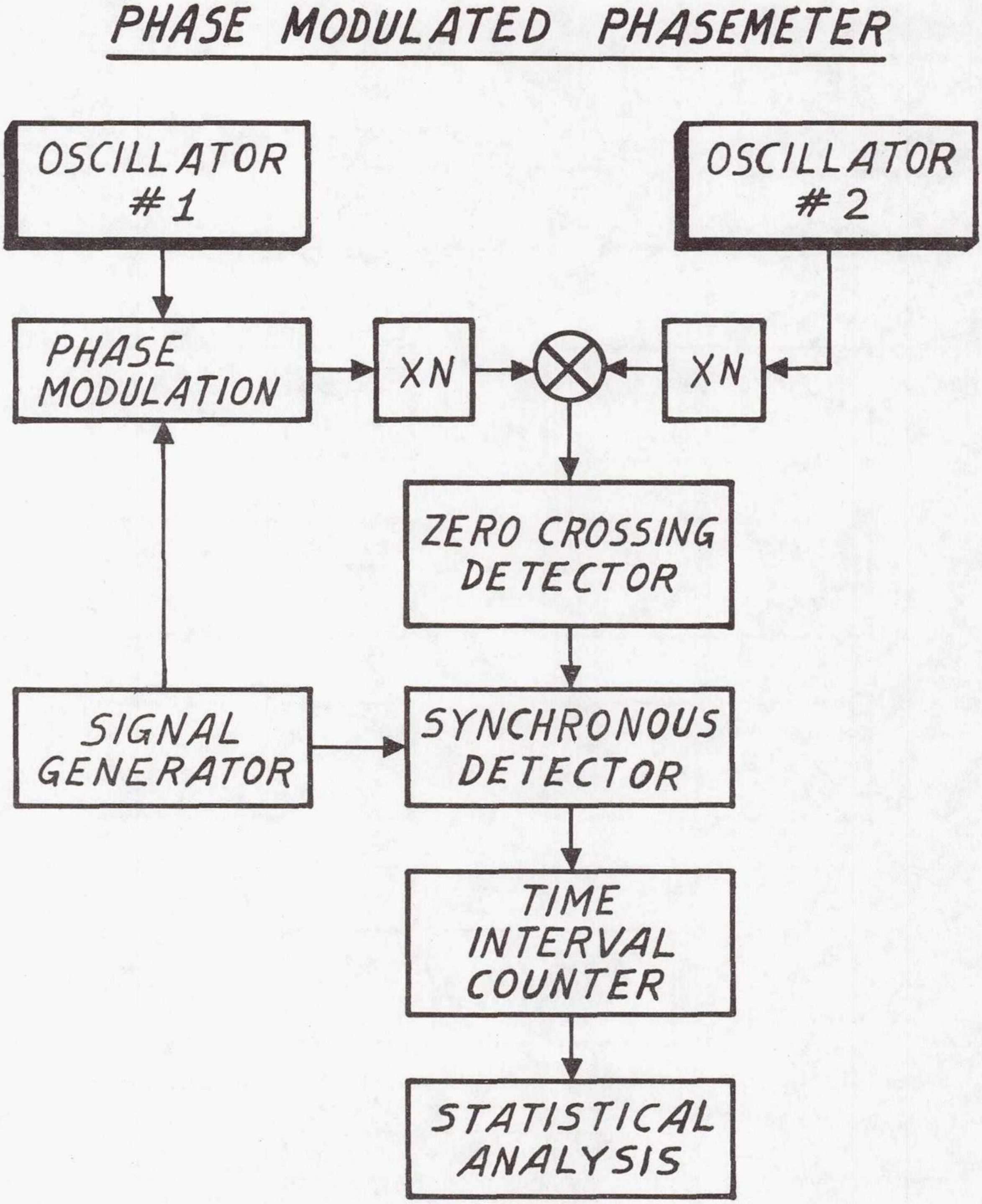


Figure 14. Block diagram illustrating the Phase Modulated Phasemeter System.

Table 1

Asymptotic Behaviour of Various Variances for Power Law Spectral Models and Polynomial Drifts

ASSYMPTOTIC BEHAVIOUR of VARIANCES

$$S_y(f) = h_\alpha f^\alpha$$

$S_y(f)$	$I^2(\tau)$	$\sigma_y^2(\tau)$	$\sigma_m^2(3,\tau,\tau)$
$h_2 f^2$	τ^{-2}	τ^{-2}	τ^{-2}
$h_1 f$	$\sim\tau^{-2}$	$\sim\tau^{-2}$	$\sim\tau^{-2}$
h_0	τ^{-1}	τ^{-1}	τ^{-1}
$h_{-1} f^{-1}$	∞	τ^{0}	τ^{0}
$h_{-2} f^{-2}$	∞	τ^{1}	τ^{1}
$h_{-3} f^{-3}$	∞	∞	τ^{2}
$h_{-4} f^{-4}$	∞	∞	τ^{3}

$$y(t) = d_n t^n$$

$d_n t^n$	$I^2(\tau)$	$\sigma_y^2(\tau)$	$\sigma_m^2(3,\tau,\tau)$
$d_1 t^1$	—	τ^2	0
$d_2 t^2$	—	—	τ^4

QUESTIONS AND ANSWERS

DR. GERNOT M. R. WINKLER, U. S. Naval Observatory:

I have nothing to criticize. I think it was a very clear explanation. However, I think there is still a problem if you want to introduce a generally educated person to the subject because of the terminology, which was, of course, adopted a long time ago; it is misleading. And I would like to suggest an additional approach to such a generally educated person.

There are two ways, two main distinctions, in which we can measure and/or characterize frequency instability. The first one, called time domain measurement, essentially measures and/or characterizes the instabilities by measuring phase differences. And we obtain a statistical distribution of the carrier frequency, and we characterize that statistical distribution of the carrier frequency.

In the frequency domain, we interpret the variations and measure them as variations of the modulation frequency or of a modulation frequency around a fixed carrier.

Now, I think that this is the first thing. Now in the time domain, there are again two essentially different methods to do it. One is to obtain samples of the time readings which are equally spaced, and then you look at the statistics and have various ways to characterize it.

The other way essentially is counting phase differences between zero crossings, and you obtain your desired statistics this way. Now by doing that you inevitably have the question of whether you have dead time or not dead time. And you have the various variations of your methods.

But I believe the essential point which we tried to get across is that the distinction of time domain or frequency domain is not in frequency, but it is the distinction of frequency measurements of a carrier or frequency measurements of a modulation frequency around a fixed carrier.

In both cases, we really measure frequency. But they mean two different things. Thank you.

DR. HARRY PETERS, Sigma Tau:

I also thought that was one of the best summaries that I have ever heard on the subject. If you will allow me, I would like to make one comment regarding the use of such systems.

In the use of any of these systems, you need a frequency referance. And preferably the reference should be much better than the things you wish to measure. If you wish to measure a crystal in a rubidium in 1 to 100 seconds averaging time, of course you have no standard which is significantly better. And for such systems, the

use of a cesium as a reference in all of these systems is particularly inappropriate because you need a subsidiary standard as sort of a flywheel for all measuring times out to on the order of a thousand seconds. And this is why many people want hydrogen masers in their systems. For any of these systems, you would like a standard which is superior to all the other standards for all the measuring times in which you are interested. I had many other extensions of these comments, but I think I will stop here. Thank you.

MR. DAVE ALLAN, National Bureau of Standards:

Let me clarify one thing the novice to the community. I think they have done an outstanding job in showing how you can characterize an oscillator as to the random fluctuations basically. One must be very careful; to really characterize an oscillator, there are all kinds of systematic effects that must be included as far as the manufacturer and the user are concerned. The dependence upon temperature, pressure, humidity, whatever you have, that the oscillator may depend upon, is a whole set that must be included in a proper characterization of an oscillator.

Page intentionally left blank

A Modular Multiple Use System for Precise Time and Frequency Measurement and Distribution

by

Victor S. Reinhardt
William A. Adams
Grace M. Lee
NASA-Goddard Space Flight Center

and

Robert L. Bush
Bendix Field Engineering Corporation

ABSTRACT

A modular CAMAC based system is described which has been developed at NASA/Goddard Space Flight Center to meet a variety of precise time and frequency measurement and distribution needs. The system is based on a generalization of the dual mixer concept. By using a 16 channel 100 ns event clock, the system can intercompare the phase of 16 frequency standards with sub-picosecond resolution. The system has a noise floor of 26 fs and a long term stability on the order of 1 ps or better. The system also uses a digitally controlled crystal oscillator in a control loop to provide an offsettable 5 MHz output with sub-picosecond phase tracking capability. A detailed description of the system is given including theory of operation and performance. Also discussed is a method to improve the performance of the dual mixer technique when phase balancing of the two input ports cannot be accomplished.

INTRODUCTION

This paper describes a modular CAMAC based distribution and measurement system developed at NASA/Goddard Space Flight Center to meet a variety of precise time and frequency needs. The system grew out of the desire to minimize the development effort and costs for three system requirements at Goddard Space Flight Center:

1) a frequency combiner/selector (FCS) capable of providing a

reliable 5 MHz output tracked to the average phase of multiple atomic frequency standards, but on time with respect to UTC.

2) an automated data acquisition system capable of intercomparing the phase of multiple atomic frequency standards with sub-picosecond resolution and capable of correlating these intercomparisons with environmental data and UTC.

3) a remote measurement and distribution system (RMDS) capable of distribtuing the 5 MHz output of a frequency standard to a remote site, without degradation in phase stability, for comparisons against remote frequency standards.

System Description

A block diagram of the generalized measurement and distribution system is shown in figure 1. The heart of the system is a phase comparison system which works on a generalization of the dual mixer technique.[1] The phase comparison system includes the multiple mixers, the offset crystal oscillator and the 100 ns multichannel event clock shown on the block diagram. Not shown on the diagram, but essential for proper system operation, are buffer amplifiers which isolate the various 5 MHz channels from each other and low noise zero crossing detectors at the output of each mixer. This phase comparison system uses the offset crystal oscillator to heterodyne into 1 Hz beats 5 MHz inputs from sources whose phases are to be monitored. For the purpose of maintaining system operation independent of any of the 5 MHz inputs, the offset crystal oscillator is left free running. The multichannel event clock records the epoch of these beats to 100 ns. As shown in the theory section, this effectively records the differences in phase between each 5 MHz input and the crystal oscillator to a resolution of 0.02ps. By taking the difference in epoch between two channels, the data bus controller can determine the difference in phase between any two 5 MHz inputs with this same resolution. The multichannel event clock is also used to monitor 1 pps or other TTL pulse sources directly with 100 ns resolution.

The distribution functions of the system are implemented by the D/A controlled crystal oscillator (DXCO) and the RF output control shown on the diagram. Again, not shown but essential to proper operation, are buffer amplifiers in the 5 MHz lines. The frequency of the crystal oscillator is controlled via an 18 bit D/A converter with a fractional resolution of 7.6×10^{-13} per bit. Since the DXCO output is fed back to the phase comparison system, the data bus controller can use the D/A converter to control the phase of the DXCO with sub-picosecond resolution. Actual phase performance depends on the phase noise of the other 5 MHz inputs, the phase noise of the crystal oscillator, and the time constant of the control loop. The RF output control is essentially

an RF switch which connects the 5 MHz output to a back up 5 MHz source in case of control system failure.

Data transfer and logical functions are implemented by a data bus controller. This unit can be a microprocessor built in to the system or an external processor. Of course, data output and program control devices are also part of the system for communication to and from the processor.

HARDWARE DESCRIPTION

The data bus chosen was the IEEE and IEC instrumentation standard called CAMAC.[2] A CAMAC crate with several of the instrumentation modules is shown in figure 2. The advantages of using CAMAC are many fold. First, since CAMAC is an IEEE and IEC standard, one can interchange hardware from many manufacturers. Existing instrumentation can be used with several stand alone microprocessor based bus controllers or can be interfaced to virtually any minicomputer through interface type controllers. Second, many off the shelf modules are available to accomplish PTTI functions. For example, should the need arise to extend the resolution of the event clock, an IEEE 488 (HPIB) interface is available so high resolution time interval counters can be used. Alternatively, time interpolators with resolutions of 50 ps or better are available which can be used in conjunction with the multichannel event clock. Finally, the maximum data rate of the bus, 24 Mbps, is more than adequate for most purposes.

The multichannel event clock is shown in figure 3. This clock is capable of recording the epoch of 16 channels of TTL events (positive or negative edge selectable for each channel) to 100 ns. Whenever one or more events occur during a 100 ns interval, the epoch is recorded in a FIFO memory along with a channel identity word. In this way overlap and dead time problems are taken care of. The FIFO memory can be read as needed by the processor as long as the capacity of the FIFO is not exceeded. Epoch is stored to one day. When one day is exceeded an interrupt is sent to the processor and a bit is set until the processor resets it. In this way, the processor can keep track of the day and year.

The heart of the system is the analog part of the phase comparison system. A brass board is shown in figure 4. This system is also modular. The four larger modules are buffered mixers and zero crossing detectors. The smallest modules are buffer amplifiers and the intermediate module is an eight channel driver amplifier. These units exhibit very low phase noise (see theory section) and low environmental coefficients. Figures 5 and 6 show some typical temperature coefficient measurements of a driver amplifier and a buffered mixer. The buffer amplifiers are essentially identical to the driver amplifier. All temperature coefficients have been measured at 1ps/$^{\circ}$C or less.

Figure 7 shows voltage coefficient measurements. Notice that all voltage coefficients are approximately 1ps/V or less.

THEORY AND PERFORMANCE

The phase comparison system is a generalization of the dual mixer phase measurement technique shown in figure 8. In this technique, a transfer oscillator of nominal angular frequency $\omega_o - \varepsilon$ is used to heterodyne the outputs of two reference oscillators of nominal angular frequency ω_o to two beats of nominal angular frequency ε. As shown in the figure, a time interval counter, set up to measure the time difference between the zero crossings of the two beats, effectively measures the difference in normalized phase ($x = \phi/\omega$) between the two reference oscillators multiplied by the factor ω_o/ε. One thing not shown on the diagram but essential for cancellation of the phase noise of the transfer oscillator is that the time interval between zero crossings be significantly shorter than the correlation time of the low pass filters which determine the noise bandwidth of the system.

In the phase comparison system described here, the offset crystal oscillator is the transfer oscillator. Since N reference oscillator inputs are involved, a single N channel epoch clock replaces N-1 time interval counters. This not only reduces the required hardware, as will be shown later, this improves system operation under certain operating conditions.

To understand the detailed operation of the phase comparison system, consider a 5 MHz input on channel i of the form:

$$V_i = A_i \sin(\omega_o t + \phi_i(t))$$

and a signal from the transfer oscillator of the form:

$$V_T = A_T \sin(\omega_o t + \phi_T(t))$$

where all the phase deviation from an ideal signal of angular frequency ω_o has been put into ϕ_i and ϕ_T respectively. The mixer at channel i outputs a signal of the form:

$$V_M = A_M f(\phi_T - \phi i)$$

where f(x) is a periodic sine like function whose only important property is that $f(x) = 0$ at $x = n\pi$ (n any integer). The ith zero crossing detector outputs a positive going pulse which is recorded as an event by the 100 ns clock at time t_i given by:

$$\phi_T(t_i) - \phi_i(t_i) = 2\pi n_i$$

The difference in normalized phase ($x = \phi/\omega_o$) between channel i and channel j, for $n_i = n_j$,

is, then:

$$Dx\ (i,j) = x_T(t_i) - x_T(t_j)$$

where:

$$Dx = (\phi_i(t_i) - \phi_j(t_j))/\omega_o$$

and:

$$x_T = \phi_T/\omega_o$$

But:

$$x_T(t_i) - x_T(t_j) = \bar{y}\ Dt$$

where:

$$Dt\ (i,j) = t_i - t_j$$

and $\bar{y}$ is the fractional frequency offset of the transfer oscillator from ω_o averaged over time Dt:

$$\bar{y} = \int_{t_i}^{t_j} \frac{\omega_T - \omega_o}{\omega_o}\, dt$$

This yields:

$$Dx = \bar{y}\ Dt \qquad (1)$$

which states that the difference in the phase of any two channels is given by the difference in zero crossing times of the beats times the fractional frequency offset of the transfer oscillator.

The stability of the measurement system over the time T can be characterized by a two sample variance of Dx:

$$\sigma^2_{Dx}\ (2,\ T,\ Dt) = \tfrac{1}{2} \langle (Dx\ (t) - Dx\ (t + T))^2 \rangle$$

Using (1), this becomes:

$$\sigma_{Dx} = Dt\ \sigma_y\ (2,\ T,\ Dt) \qquad (2)$$

or:

$$\sigma_{Dx} = \sigma_x(2,T,Dt) \qquad (3)$$

where σ_y is the two sample variance for the fractional frequency variations of the transfer oscillator and:

$$\sigma_x = Dt\ \sigma_y$$

From Equations (2) or (3), the stability properties of Dx can be examined. For Dt less than, t_c, the time constant of the zero crossing detectors which determine the noise bandwidth of the system, σ_y is approximately a constant. This means that, for $Dt < t_c$; σ_{Dx} gets smaller as Dt gets smaller. In actual practice, this is limited by noise introduced by the measurement system itself. In the system described (t_c = 13 ms, f_c = 12 Hz):

$$\sigma_{Dx}\ (2,\ 1s,\ 120\mu s) = 2.7\text{x}10^{-14}s$$

Figure 9 shows the long term phase stability. Notice it is on the order of 1 ps or better for T up to days. Notice also that during this time the transfer oscillator has varied in frequency by almost 10^{-10}.

For characterizing the resolution of the phase comparison system when measuring frequency, one can form the statistic:

$$\sigma^2_{MY}\ (T) = \frac{1}{4N} \sum_{i=1}^{N} (Dx\ (iT) - 2Dx\ ((i+1)\ T) + Dx\ ((i=2)\ T))^2$$

which in the limit of large N becomes the apparent value $\sigma_y(T)$ for one channel when ideal frequency standards are input to all the channels. For the data shown in figure 9, σ_{Dx} and σ_{MY} are computed in the following chart:

System Performance
(system bandwidth = 12 Hz)

T (s)	N	σ_{Dx} (10^{-14}s)	σ_{MY}
.99	100	2.73	$3.3\text{x}10^{-14}$
9.9	100	3.41	$3.9\text{x}10^{-15}$
99	1498	10.4	$7.4\text{x}10^{-16}$
990	148	28.0	$3.1\text{x}10^{-16}$
9900	13	71.2	$5.2\text{x}10^{-17}$
10^5	(estimated)	100	$7\text{x}10^{-18}$

The value at 10^5 seconds is estimated on the basis of a few days observed performance. A graph of σ_{MY} is shown in figure 10.

For Dt greater than t_C, one can see from (2) or (3) that the phase noise of the transfer oscillator becomes important. There are two components associated with this phase noise, a short term component associated with Dt and a long term component associated with T. That is, for T > > Dt:

$$\sigma^2_{Dx} \cong Dt^2 \; (\sigma^2_y \; (Dt) + \sigma^2_y \; (T))$$

Figure 11 shows this behavior for the case where Dt is about equal to the 1 Hz offset of the transfer oscillator. In figure 11, the "raw phase" is Dt times the average value of y over the time of the plot. Notice that the raw phase has a short term noise component greater than the "phase" of figure 10. Notice also that, for this worst case, the phase tracks the crystal transfer oscillator with about 1 ps for every part in 10^{12} change in frequency.

There is a method, however, for taking out the long term component due to changes in the transfer oscillator. Since the total epoch of each channel is recorded, the frequency of the crystal relative to any channel can be obtained at any time by taking the difference in epoch of the zero crossings of that channel. To obtain the best estimate of the crystal frequency, one can use the channel with the lowest noise input, the average of many channels, or the average over many zero crossings. Using this estimate of the transfer oscillator frequency, $\bar{y}'$, a good estimate of Dx can be obtained:

$$Dx' = \bar{y}' \; Dt \qquad (4)$$

The error in this estimate is give by the difference between (1) and (4):

$$Dx - Dx' = Dt \; (\bar{y} - \bar{y}') \qquad (5)$$

By using a running estimate for $\bar{y}'$, all the long term effects of transfer oscillator changes can be taken out. This is shown in the "corrected phase" of figure 11.

SYSTEM APPLICATIONS

The most stringent application for the measurement and distribution system is the Data Acquisition Facility shown in figure 12. This facility's principle task is to intercompare many hydrogen masers with subpicosecond resolution, to monitor their phase against UTC, and to measure the effects of environmental factors on these masers. For this purpose the CAMAC module is interfaced to a PDP11 computer. Com-

parisons with UTC are made via TV and LORAN-C.

Another application is the Frequency Combiner/Selector (FCS) shown in figure 13. The FCS will use the phase comparison system and DXCO to track a 5 MHz signal to the average phase of multiple cesium frequency standards with a frequency offset. The frequency offset will be adjusted in a secondary phase lock loop to keep the 5 MHz on time with respect to UTC measured via NASA's own Tracking Data Relay Satellite System (TDRSS) or LORAN-C. The processor will carry out fault analysis on all parts of the system and modify system behavior to minimize the consequence of any fault.

A third application is a remote distribution and measurement system. As envisioned now, this system will monitor a 5 MHz signal sent from a main site down a cable and return the same 5MHz to the main site via a parallel cable. By monitoring the changes in phase down both cables, the main site will be able to determine a correction to bring the remote 5 MHz in phase with the 5 MHz at the main site. This information will be sent via modem to the remote system to control its DXCO. The main use of this system will be to intercompare remote frequency standards without loss of stability.

REFERENCES

1. D. W. Allan and H. Daams, "Picosecond Time Diference Measurement System" 29th Annual Symposium on Frequency Control (Atlantic City, 1975).

2. <u>CAMAC Instrumentation and Interface Standards</u>, IEEE Publication (1976, New York)

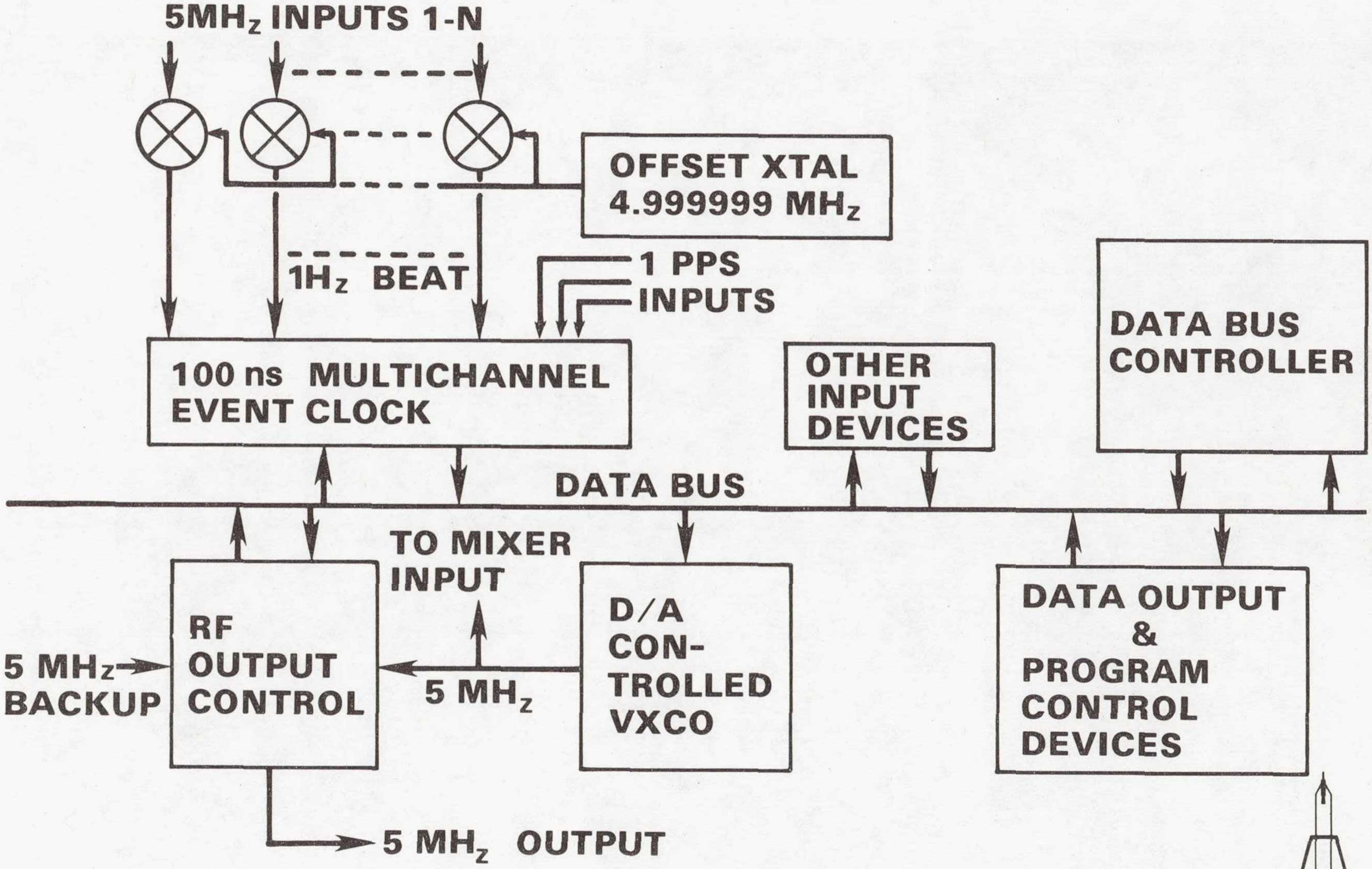


Figure 1. Distribution and Measurement System

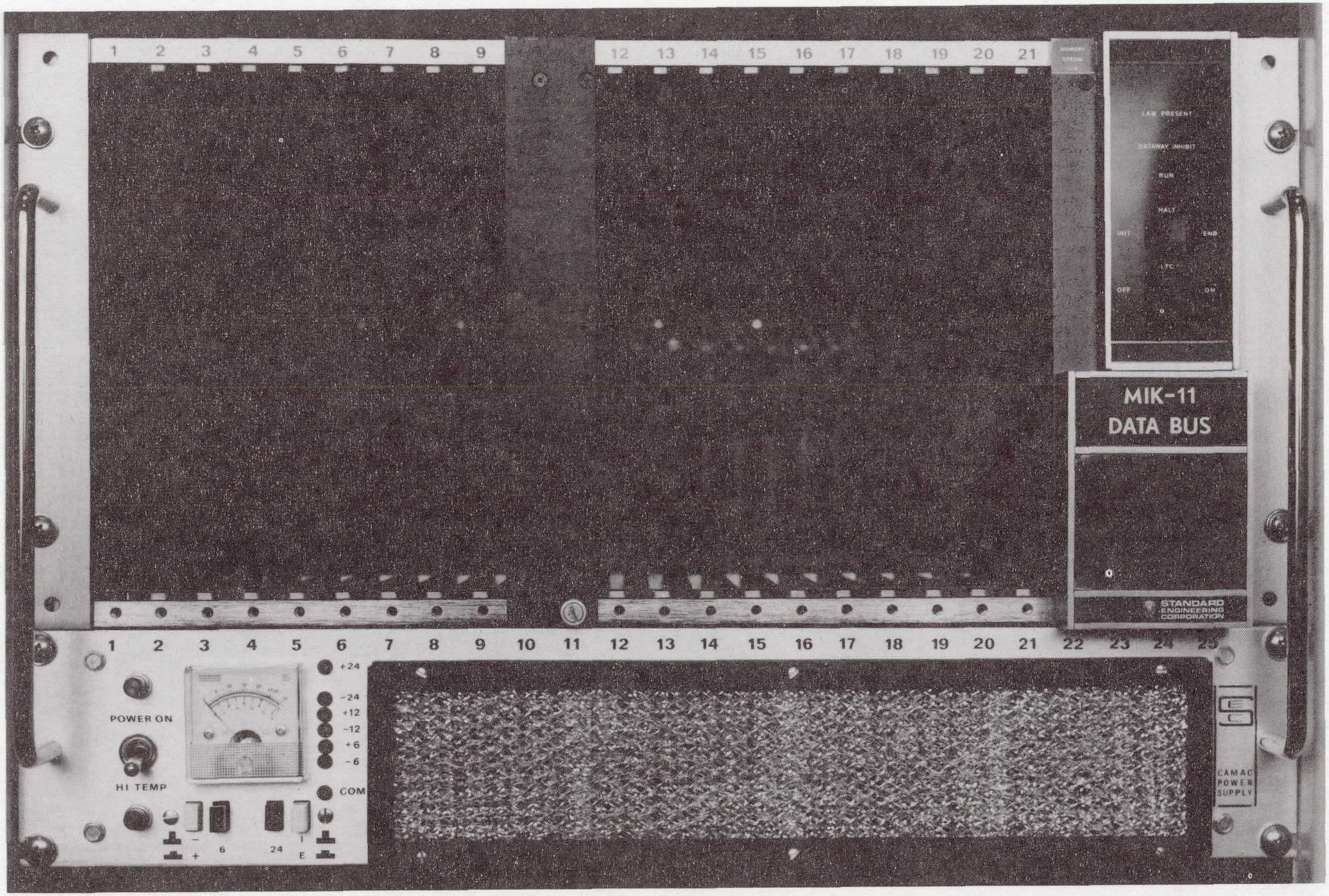


Figure 2. CAMAC Crate

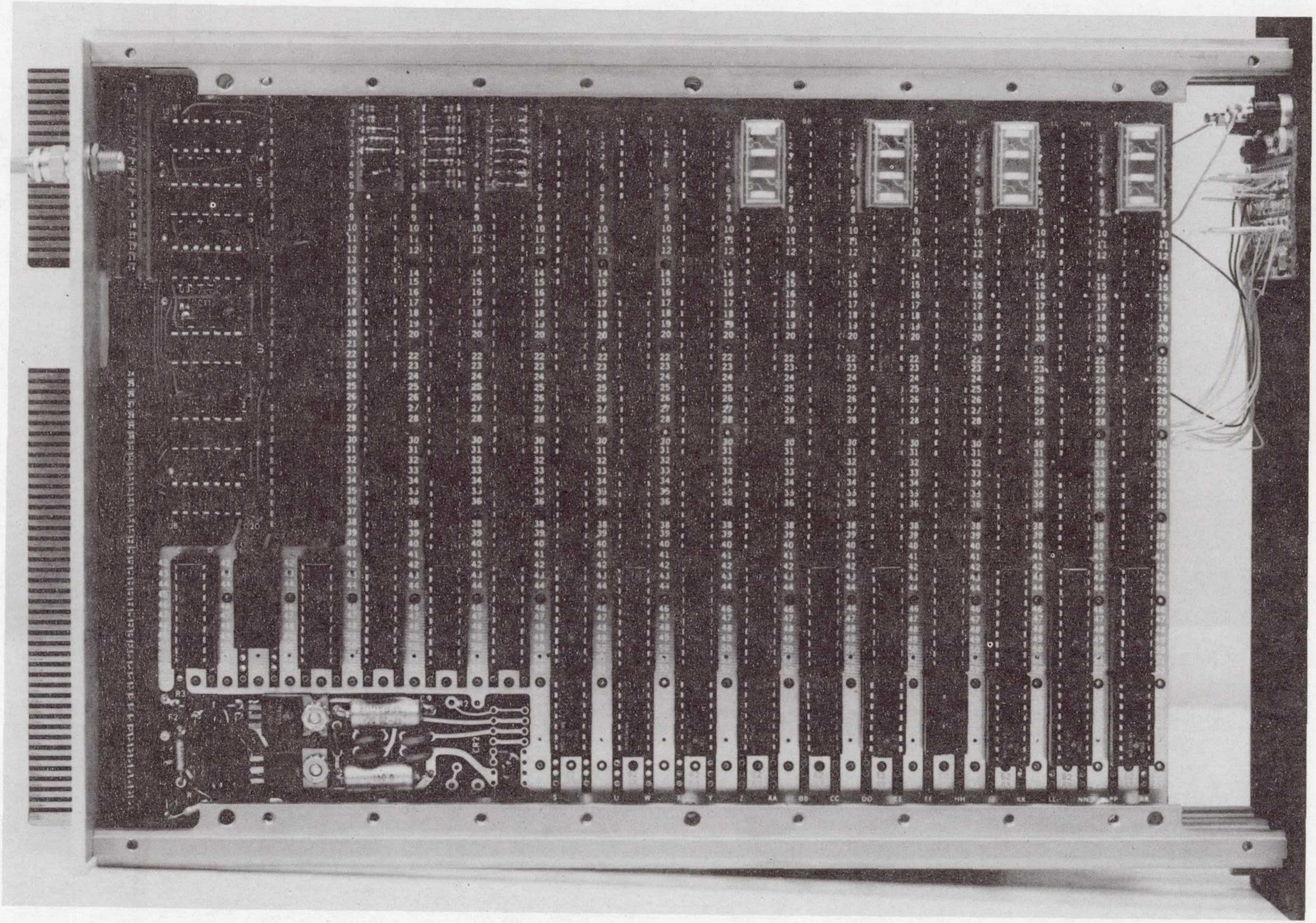


Figure 3. Real Time Clock

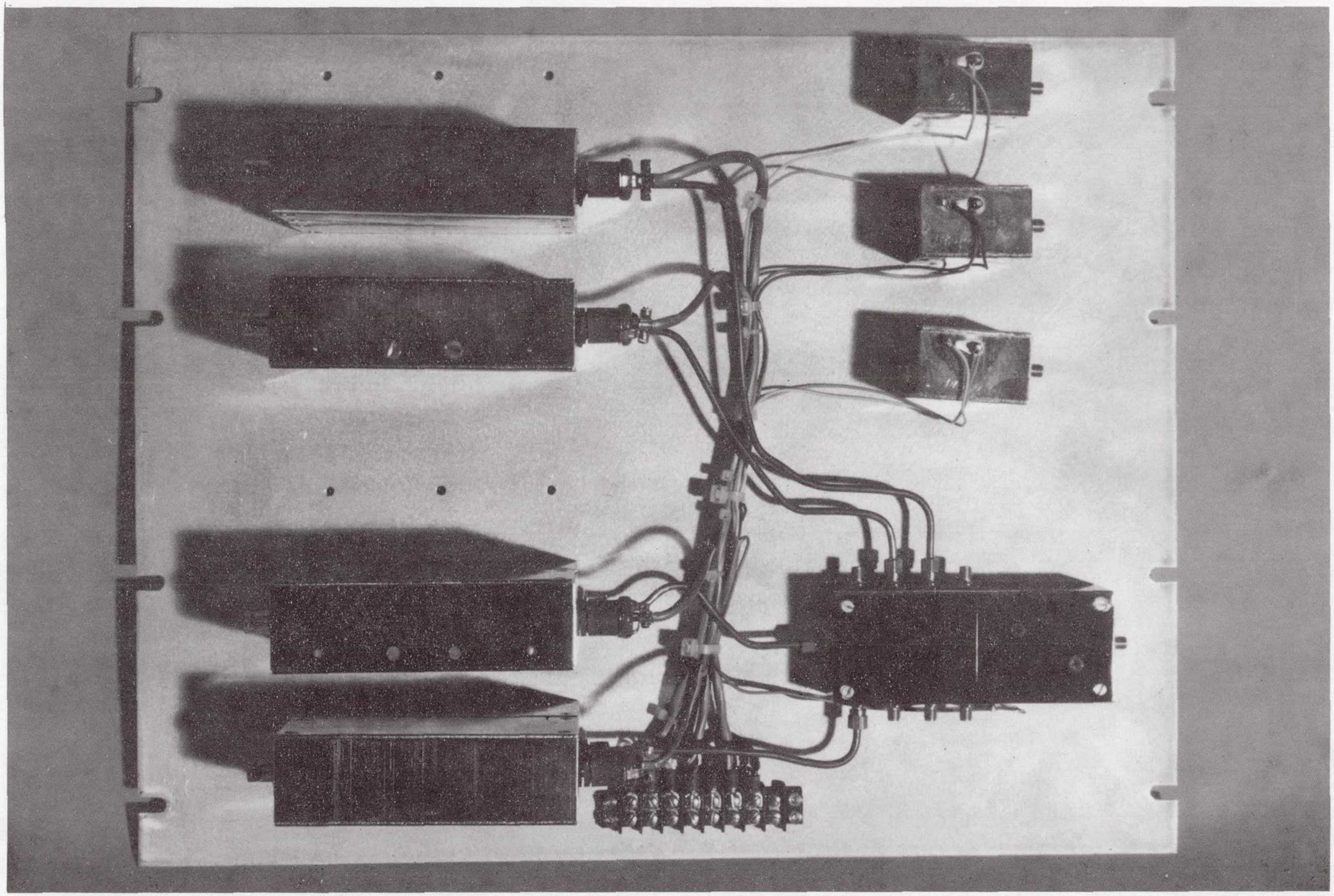


Figure 4. Phase Comparison System

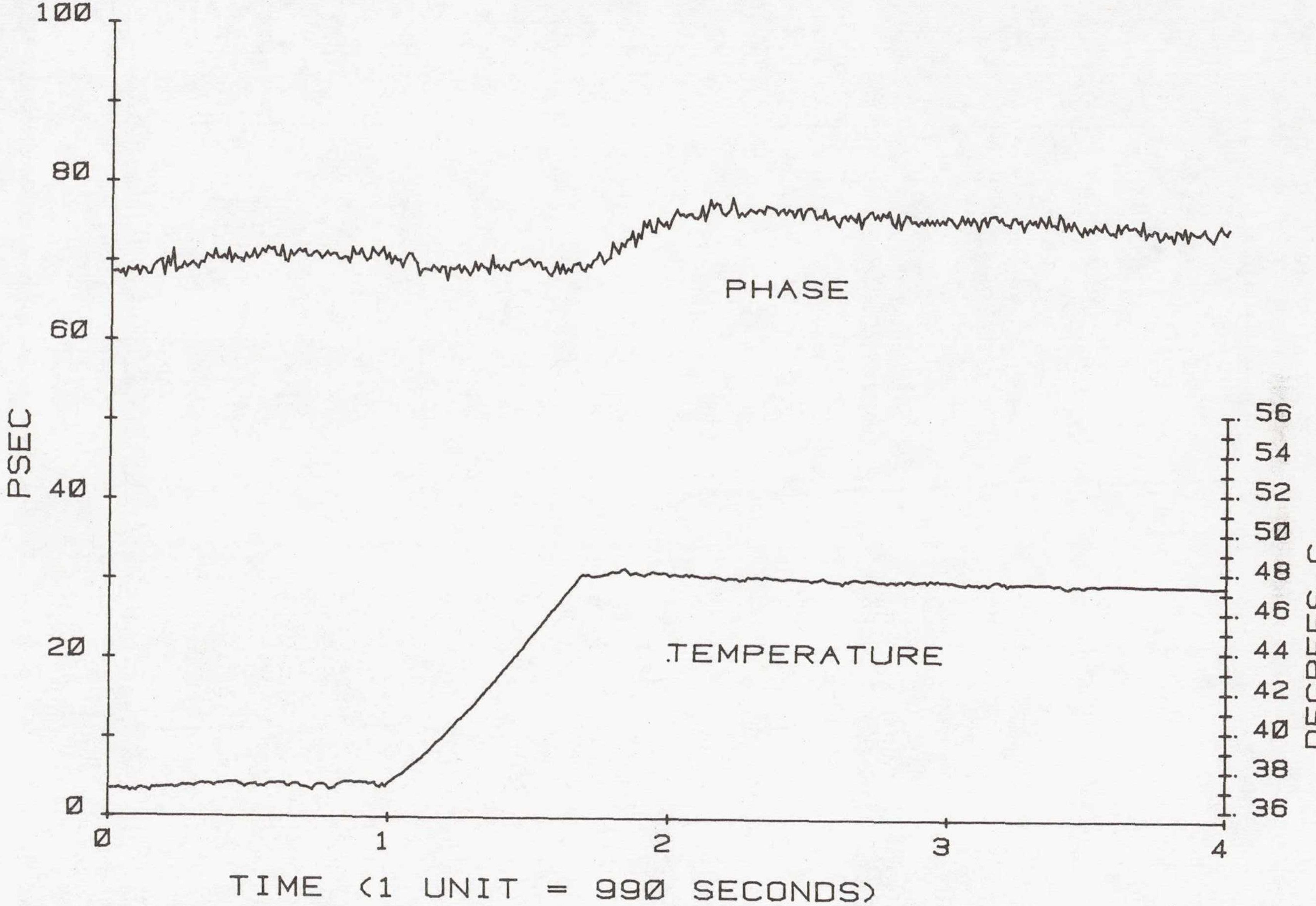


Figure 5. Temperature Coefficient of Driver Amplifier

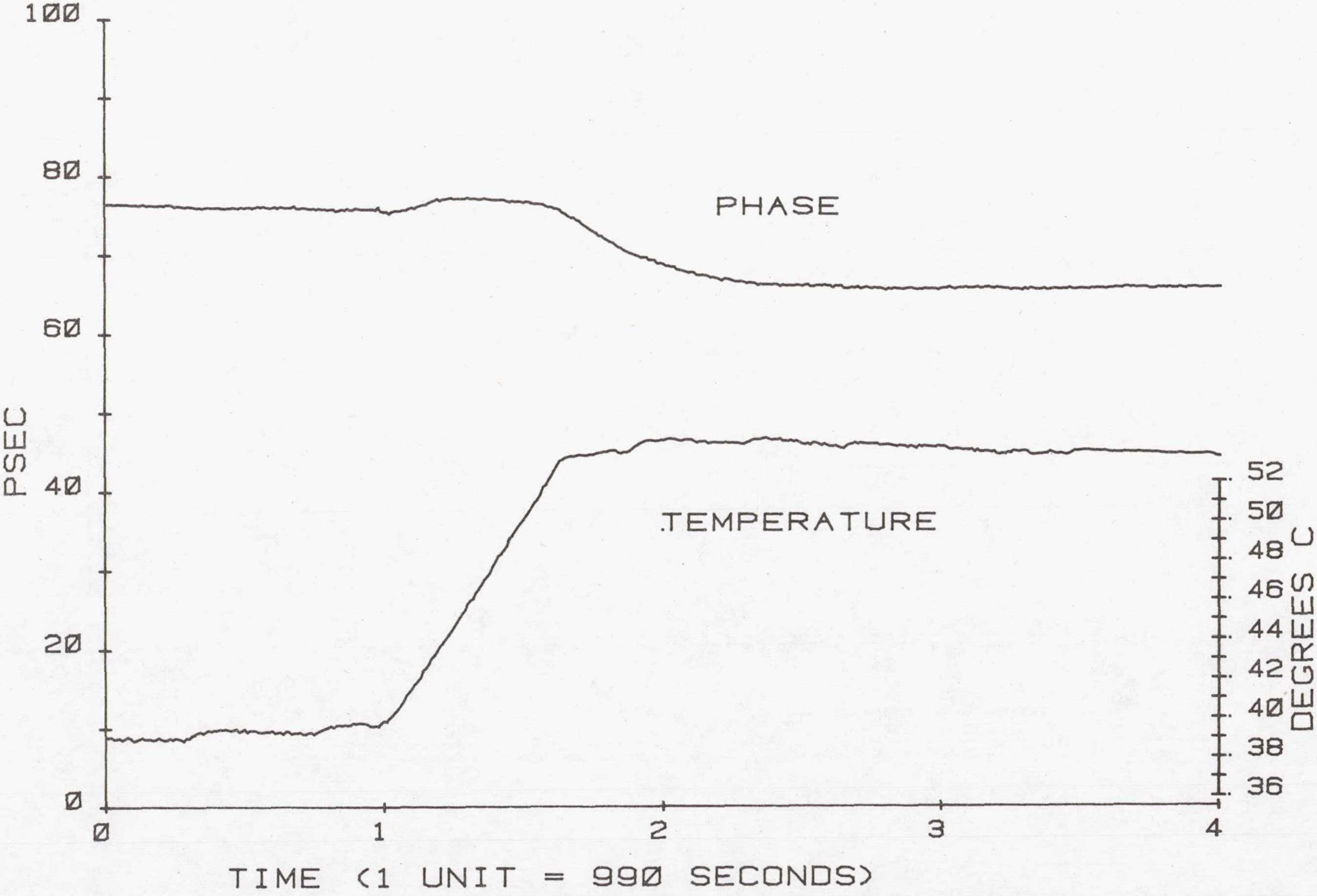


Figure 6. Temperature Coefficients of Buffer-Mixer

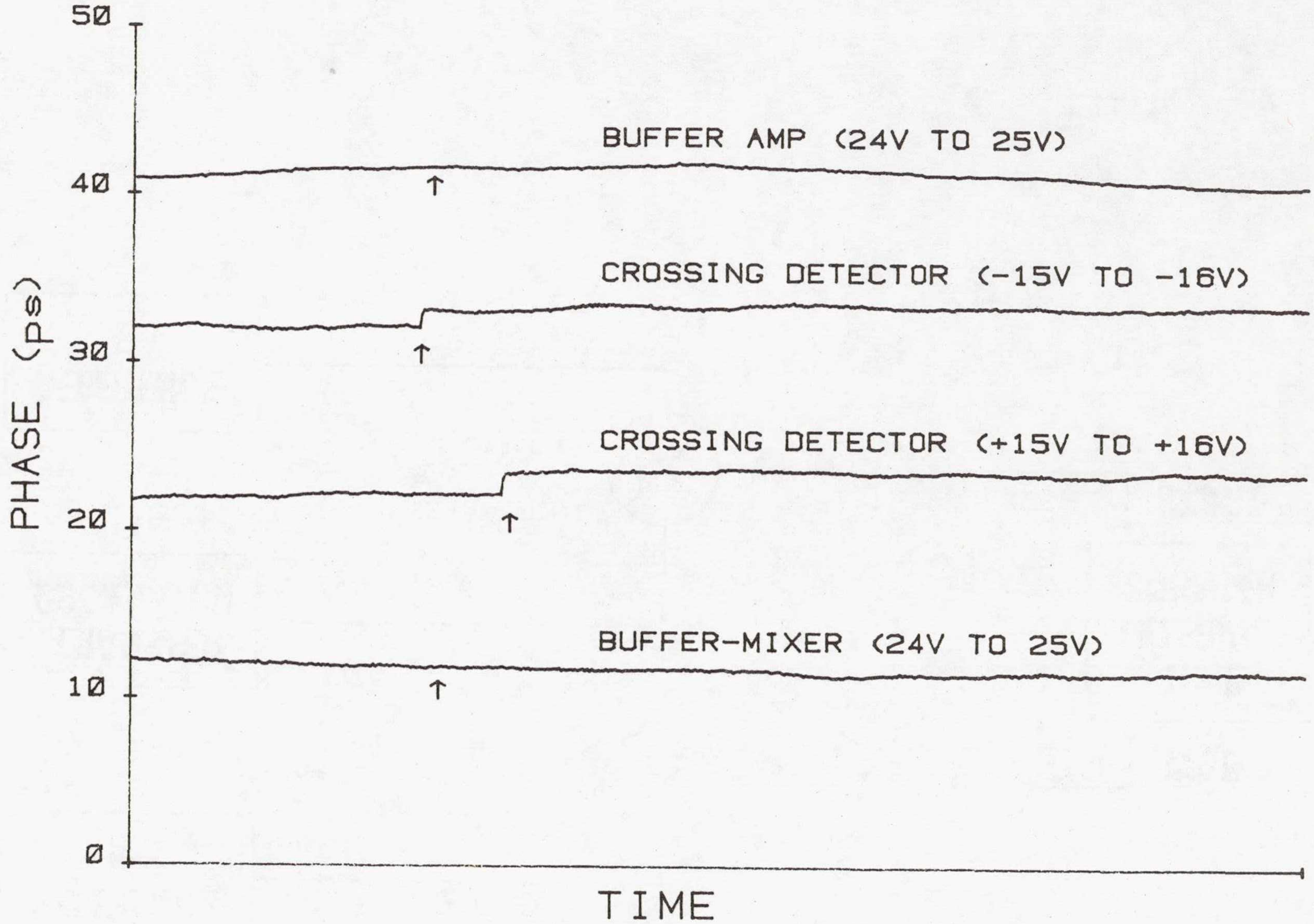
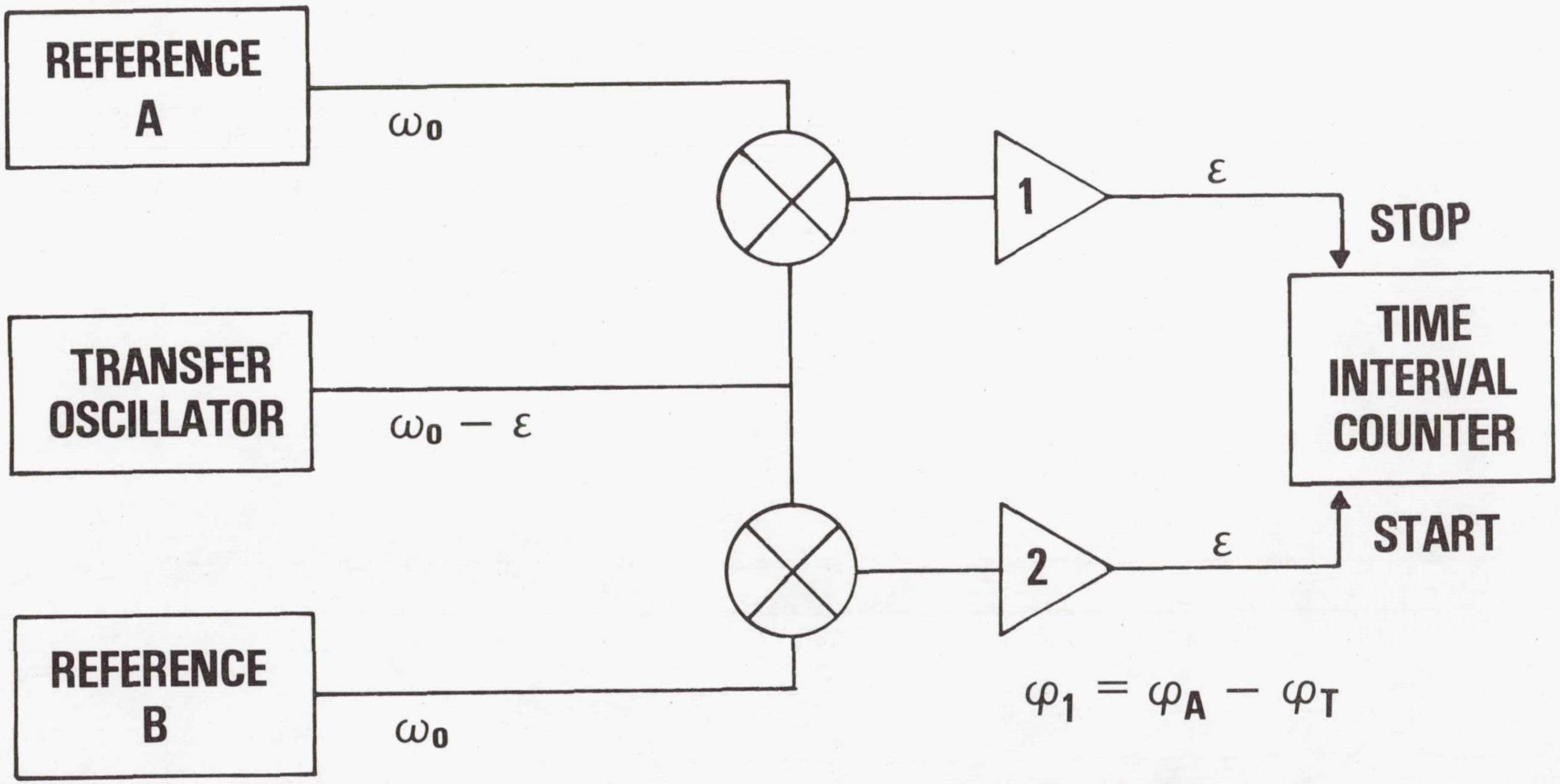


Figure 7. Voltage Coefficients



$$\varphi_1 = \varphi_A - \varphi_T$$

$$\varphi_2 = \varphi_B - \varphi_T$$

$$\varphi_1 - \varphi_2 = \varphi_A - \varphi_B$$

$$x_1 - x_2 = \frac{\omega_0}{\varepsilon} (x_A - x_B)$$

Figure 8. Dual Mixer System

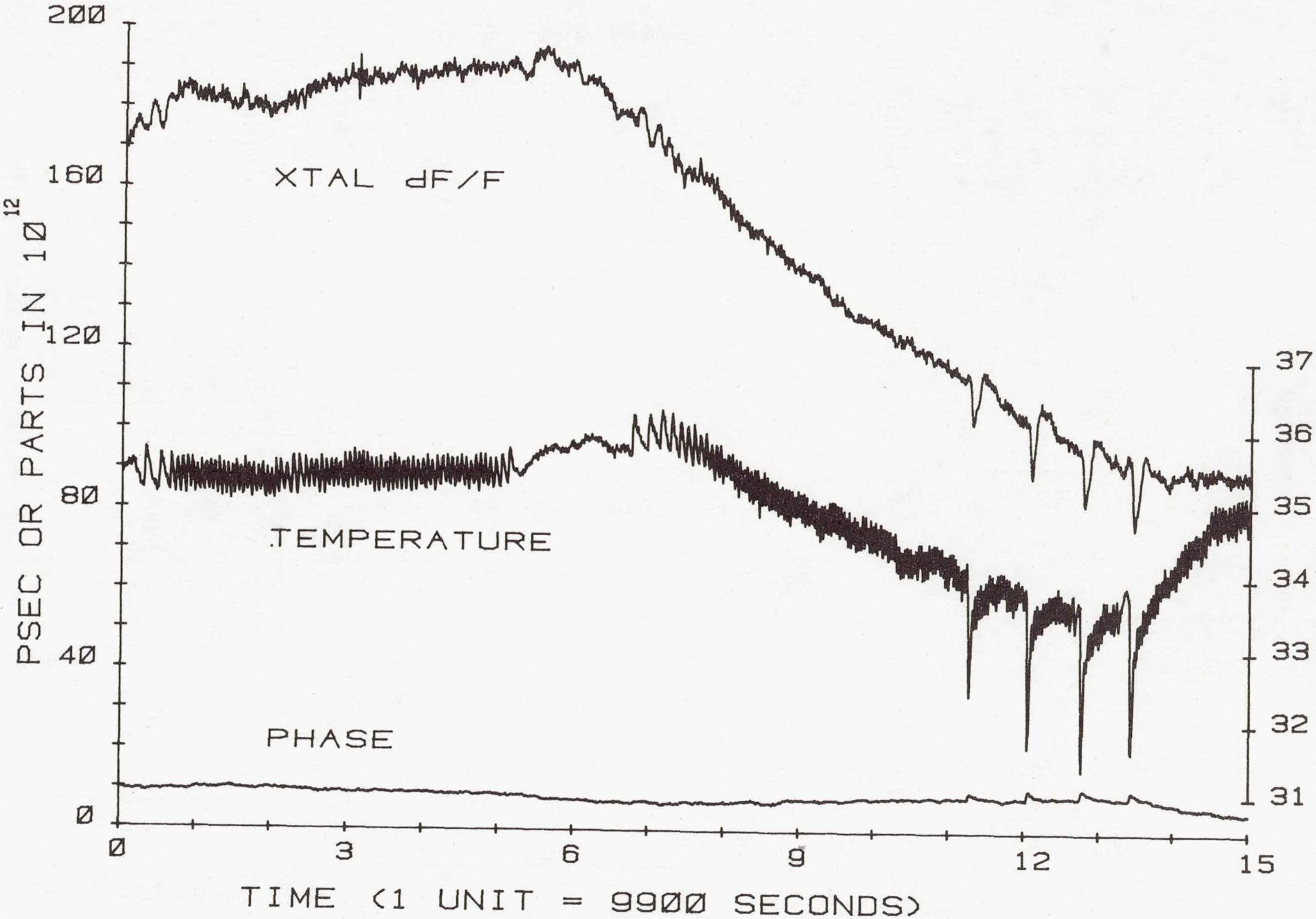


Figure 9. Phase Comparison System Performance (Dt = 120us)

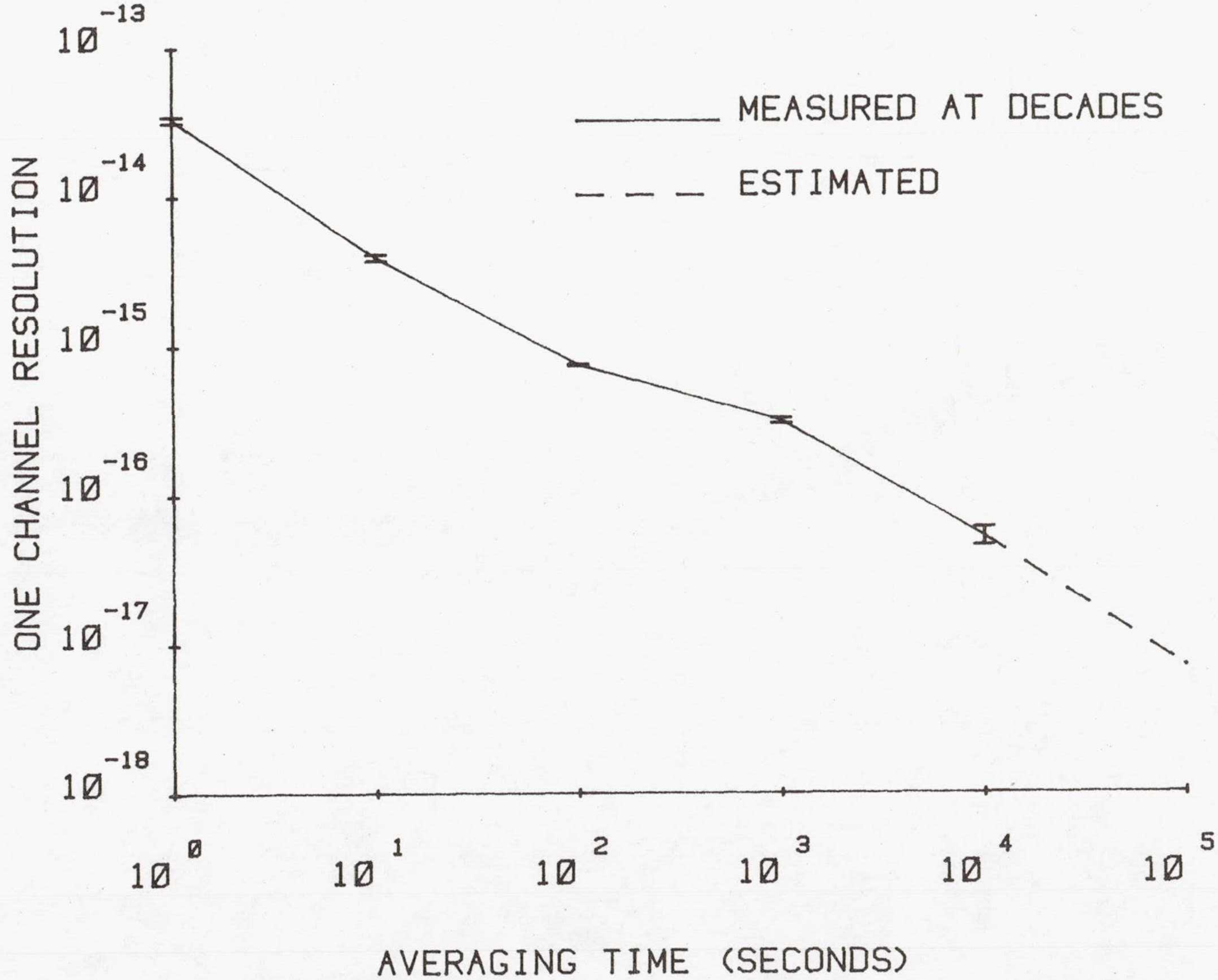


Figure 10. Frequency Measurement Results

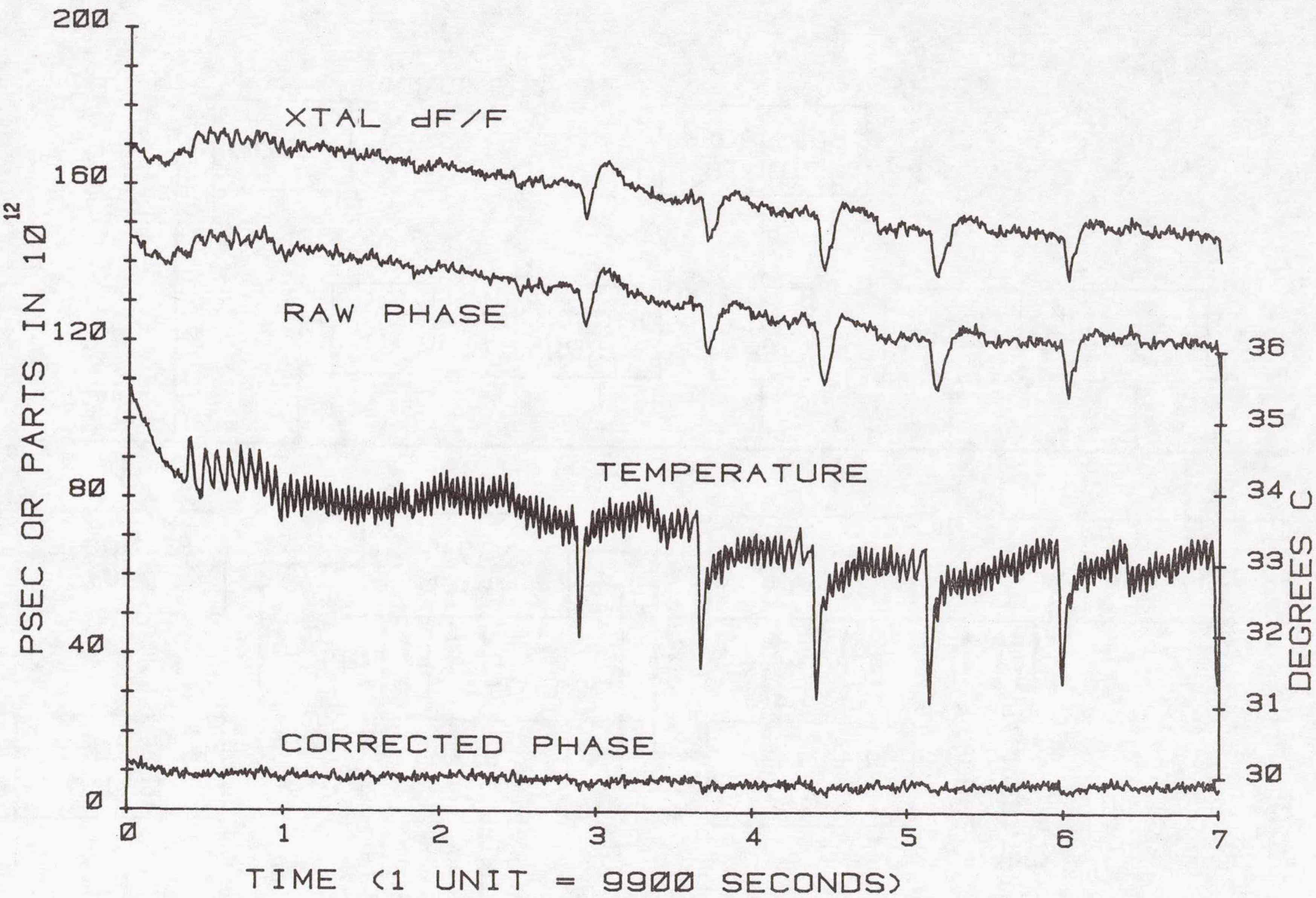
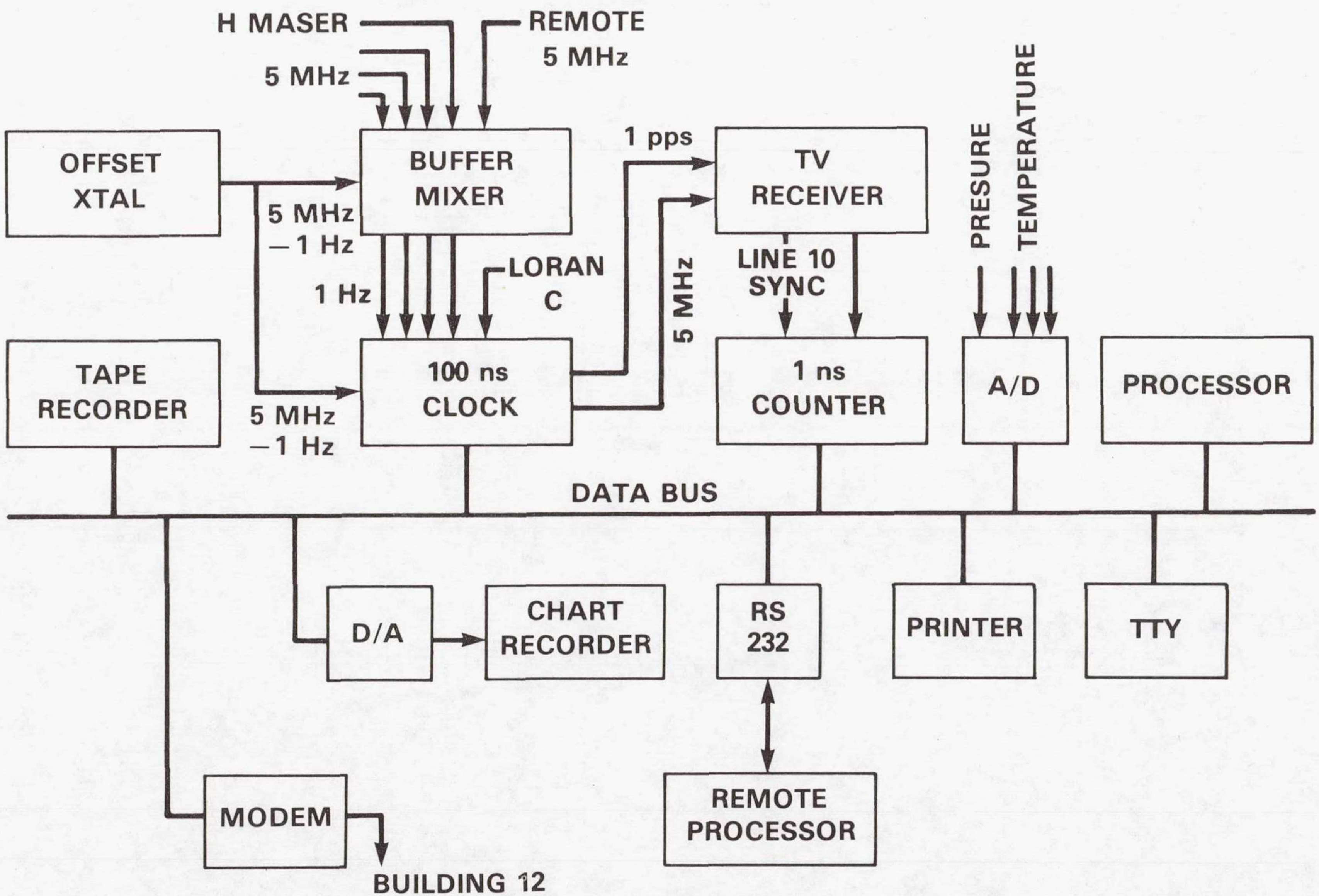


Figure 11. Phase Comparison System Performance (Dt = 990 ms)


H MASER
5 MHz
REMOTE
5 MHz
OFFSET
XTAL
BUFFER
MIXER
5 MHz
− 1 Hz
1 Hz
LORAN
C
1 pps
TV
RECEIVER
5 MHz
LINE 10
SYNC
PRESURE
TEMPERATURE
TAPE
RECORDER
100 ns
CLOCK
5 MHz
−1 Hz
1 ns
COUNTER
A/D
PROCESSOR
DATA BUS
D/A
CHART
RECORDER
RS
232
PRINTER
TTY
REMOTE
PROCESSOR
MODEM
BUILDING 12

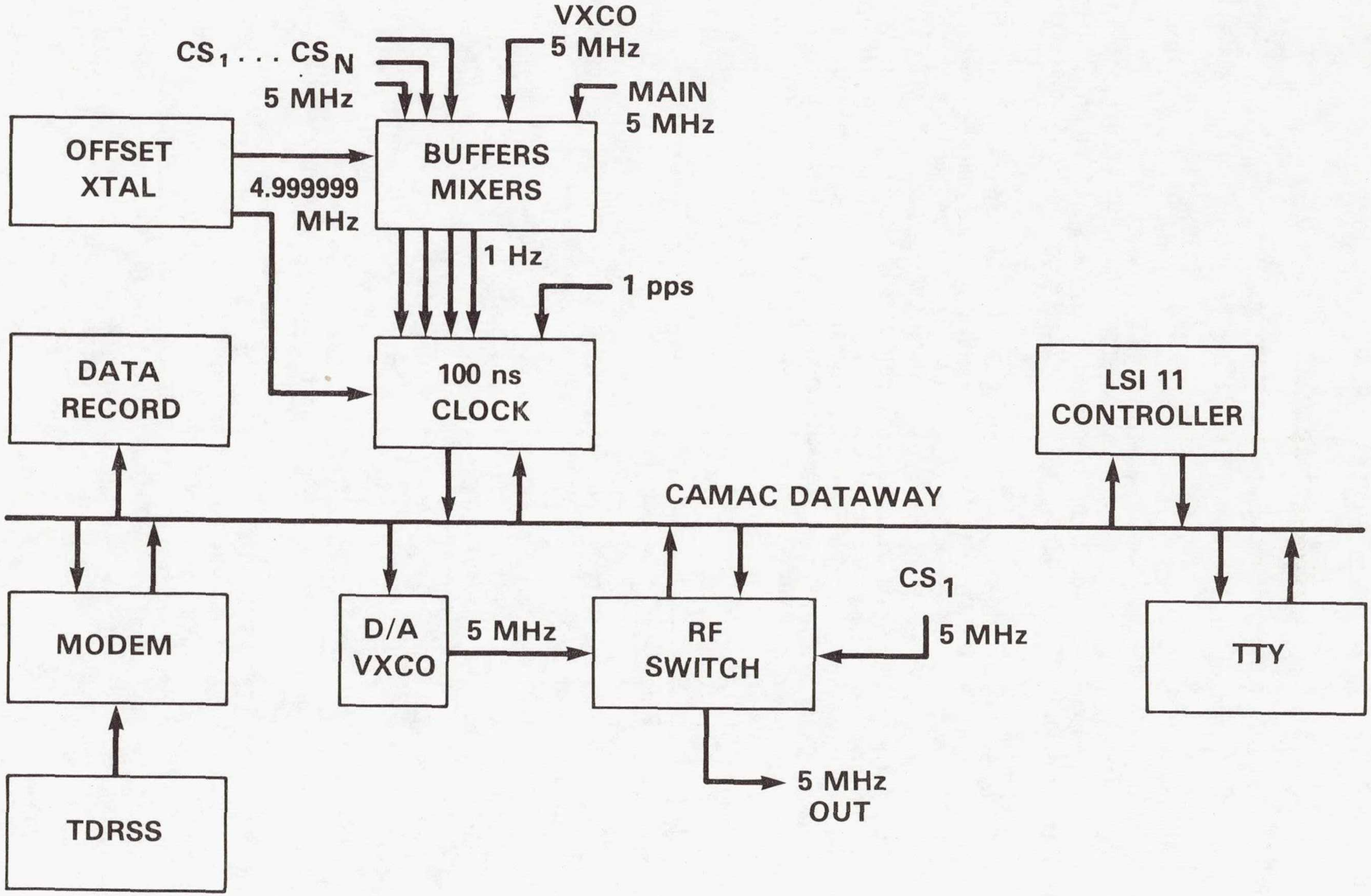


Figure 13. Frequency Combiner/Selector

QUESTIONS AND ANSWERS

DR. HELMUT HELLWIG, National Bureau of Standards:

Can the system or at least parts of it be procurred? If so, how and for how much?

DR. REINHARDT:

That is another reason we went to CAMAC. The only essential part of the system that cannot be procurred off the shelf at this time; that is, the digital parts of the system, is the 100 nanosecond clock.

One thing I didn't mention because time was short was you buy what are called klugeboards, which essentially give you all the hardware and all the interfacing to the CAMAC. And you only wire what you need. And we have wirewrap lists. In fact, if you want one of these 100 nanosecond clocks, contact us because we have been having them wirewrapped outside. It is just a matter of us sending you the drawings and the wire lists. And you can go to the same company that we did and get it wirewrapped. The price? Bob, what was the price of the 100 nanosecond clock?

SPEAKER: About $1,000...

DR. REINHARDT:

That's for the CAMAC crate itself. In fact, you can get that as a full-blown RT 11 operating system with floppy discs for about $12,000. But the microprocessors themselves cost about $3,000 or $4,000. The CAMAC crate is about $1,500.

My rule of thumb is, it is $1,000 per module--per completed module, not klugeboard. The analog parts of the system are, as you can see, also modular. We are making them in-house now. But again, if anybody wants any drawings or anything like that, they should contact us. We have printed circuit board layouts. Thank you.

DR. CARROLL ALLEY, University of Maryland:

How many cesium standards do you plan to include in your ensemble?

DR. REINHARDT:

We have the optimum ensemble for statistical analysis--two. That way, nobody can tell if anything is wrong. Right now we plan to use two. And you obviously get a factor of two improvement in reliability. The idea is to also at least divide the phase ____ by a square root of two while you are doing that. And I have an ulterior motive in this. You can plug a hydrogen maser directly into that system and have it work. It will reproduce the phase specs of a hydrogen maser.

DR. SAMUEL STEIN, National Bureau of Standards:

You can accomplish very much the same results, to remove the instabilities of a common oscillator and to gain the benefit of an event clock, in a small system in hardware, by locking your transfer oscillator to one of your references.

MR. DAVID W. ALLAN, National Bureau of Standards:

In addition to that, if you get an oscillator which has very little phase noise over the beat interval, in your case one second, the noise then becomes immaterial regardless of the phase. We have chosen as one system at NBS to use a 500 hertz beat and a very low, high-quality low-phase noise oscillator. And the phase noise just doesn't enter in.

DR. REINHARDT:

I would like to add a comment to that. The data were actually taken without the full-blown system, and now with these programmable calculators and counters, we just set the counter to alternately measure time interval and period. And so you can effectively do the same thing on a small system with a conventional dual mixer system by just changing the software.

Page intentionally left blank

A REVIEW OF THE U.S. NAVY'S PRECISE TIME AND TIME INTERVAL (PTTI) PROGRAM

RALPH T. ALLEN

NAVAL ELECTRONIC SYSTEMS COMMAND

ABSTRACT

In September 1975, the Navy's PTTI Program experienced a restructuring of its internal management. Since that time, the Program has undergone significant redirection.

This paper briefly outlines in broad terms:

(a) The history of the Navy's PTTI Program
(b) The Navy's current and projected PTTI requirements and capabilities
(c) The current and projected Program efforts

The United States Navy's Precise Time and Time Interval (PTTI) Program is a Navy-wide effort to provide Navy platforms with PTTI information traceable to the U.S. Naval Observatory (NAVOBSY).

At the Ninth Annual PTTI Applications and Planning Meeting, Rear Admiral Fowler briefly outlined the Navy's early dependency on time for navigation and the need for improved timing as ships' speeds increased with the conversion from sail to steam propulsion. He also outlined the NAVOBSY's very early involvement with timekeeping in support of this timing requirement and the consequent evolution of the NAVOBSY as the nation's timekeeper and a leader in today's international PTTI

community.

The Naval Research Laboratory (NRL) is another, more recent, example of the Navy's participation in the fields of timing and PTTI. NRL's involvement with crystal oscillator and frequency synthesizer technology, its' experiments in transferring time via the Defense Satellite Communications System (DSCS) and the development of the Hawaii Test Bed, both in conjunction with the NAVOBSY and the Naval Electronic Systems Command (NAVELEX), and its' recent efforts with atomic standard developments under the Global Positioning System (GPS) Program have resulted in numerous advances in PTTI.

In recognition of this early Navy leadership in timekeeping and timing developments, the NAVOBSY in 1956 was directed to provide the Standards of Time and Time Interval for the U.S. Department of Defense (DOD). Further, in 1965, by DOD directive, the NAVOBSY became the single DOD component responsible for PTTI management control functions.

Then, in 1972, the Chief of Naval Operations (CNO), tasked the NAVOBSY with maintaining the Time and Time Interval Standards for the Navy and assigned the Chief of Naval Material (CNM) as "The single Department of the Navy Manager responsible for PTTI."

CNM then designated the Commander, Naval Electronic Systems Command (COMNAVELEXSYSCOM) as the "Department of the Navy Manager for PTTI."

The current Navy PTTI Program and its associated program management organization, depicted in Figure 1, are the direct result of these various directives and taskings.

After establishing an official program organization and responsibilities, CNO, in September 1975, issued a draft Operational Requirement (OR) for Timing and Synchronization (PTTI). The draft OR described the operational problem as follows:

> There is a present and future need for precise time and time interval traceable to the U.S. Naval Observatory.
>
> Currently, each warfare sponsor provides his own precise time and time interval using methods which range from crystal oscillators to atomic clocks.

PRECISE TIME AND TIME INTERVAL (PTTI) PROGRAM ORGANIZATION

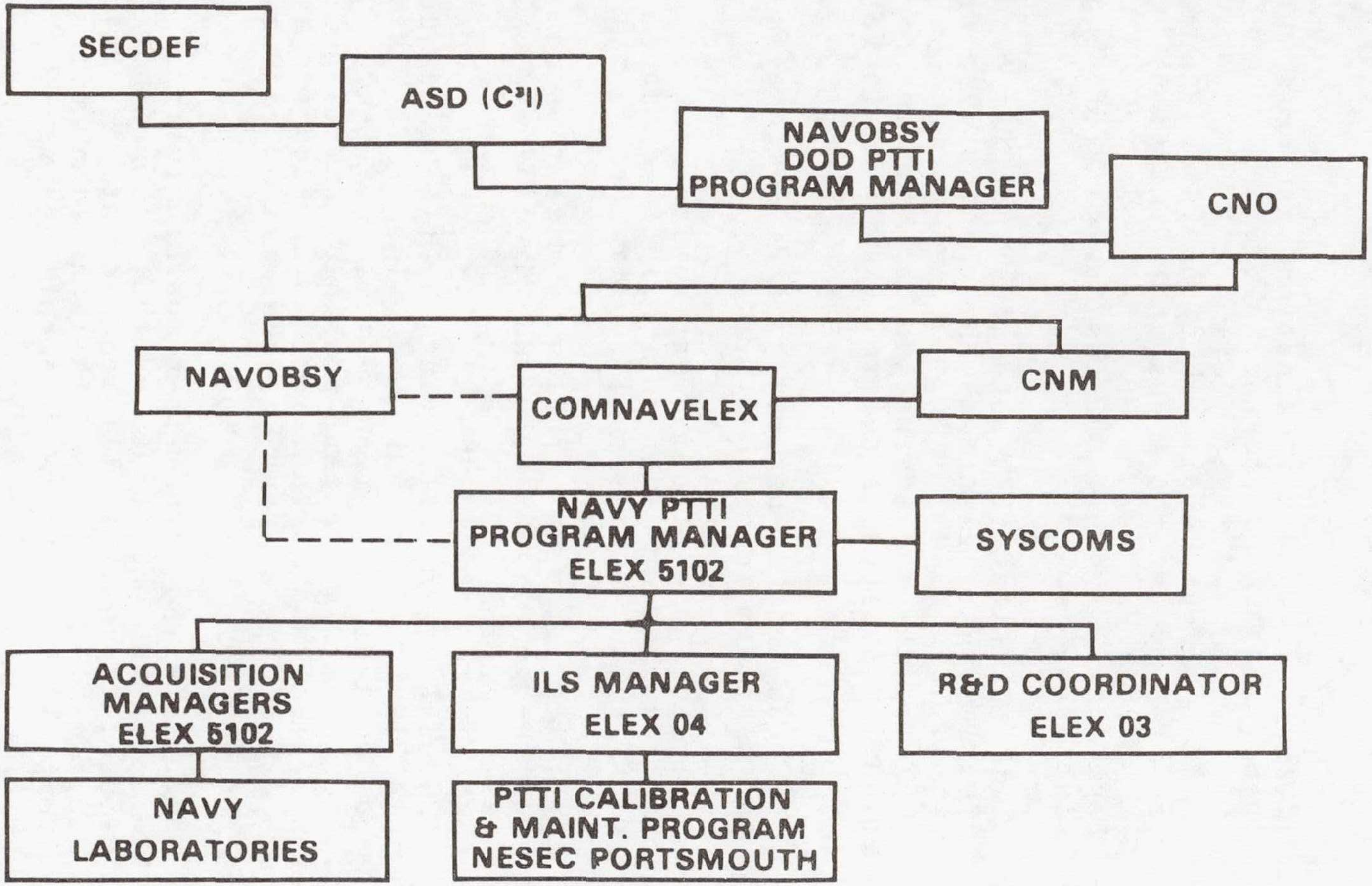


FIGURE 1

. This proliferation is neither economical nor technically sound in design and use.

The draft OR also stated that the operational requirement was for:

. The worldwide dissemination of precise time and time interval traceable to the NAVOBSY.

. Cost effective distribution systems for the various platforms to support their designated missions.

Based on that draft OR, the Navy's PTTI Program objectives are:

. To quantitatively determine the PTTI requirements of the various Navy Platforms through the 1990 time-frame.

. To review the status and capabilities of existing dissemination systems to determine if the platform requirements can be met.

. To design, as necessary, a dissemination system to meet the platform requirements.

. To design platform distribution systems to satisfy platform requirements in a cost-effective manner subject to survivability and operational constraints.

A PTTI Program Master Plan to meet these objectives, was published on 1 June 1976 and forwarded to CNO for approval.

In the Program Master Plan, a PTTI System concept, as shown in Figure 2, was outlined. That PTTI System concept is based on a dissemination system, see Figure 3, which uses, to the maximum extent possible, existing dissemination systems and a platform distribution system, see Figure 4, which also uses, to the maximum extent possible, existing equipments.

Because of an inability to obtain internal approval of the draft OR, CNO requested that NAVELEX perform an analysis of the Navy's PTTI requirements as outlined in the Program Master Plan and that the plan be resubmitted upon completion of the requirements analysis.

NAVELEX did perform the PTTI Requirements Analysis. It was completed in July 1978 and the Final Report was

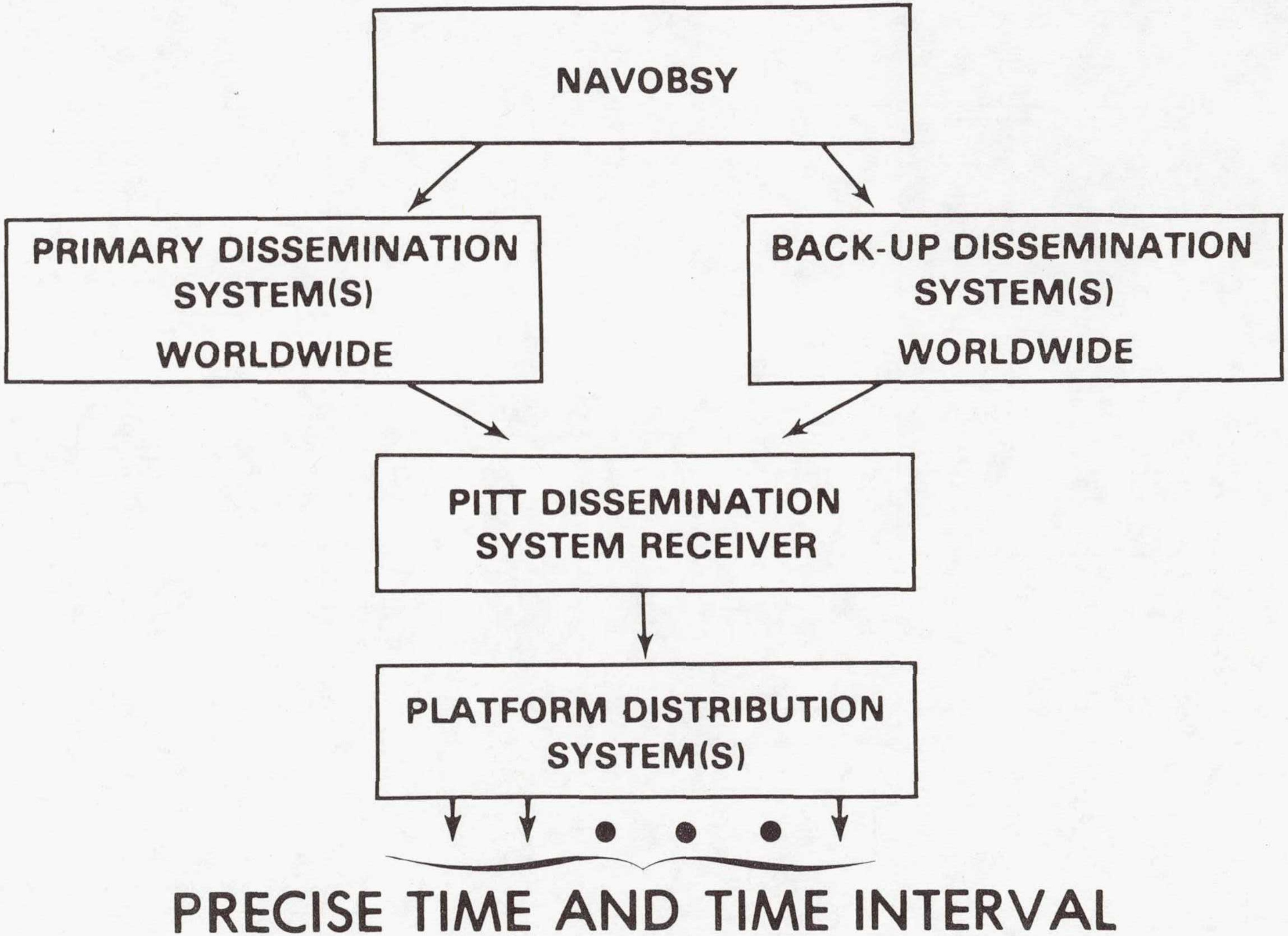

NAVOBSY
PRIMARY DISSEMINATION SYSTEM(S) WORLDWIDE
BACK-UP DISSEMINATION SYSTEM(S) WORLDWIDE
PITT DISSEMINATION SYSTEM RECEIVER
PLATFORM DISTRIBUTION SYSTEM(S)
PRECISE TIME AND TIME INTERVAL

FIGURE 2

PRECISE TIME AND TIME INTERVAL (PTTI) PROGRAM DISSEMINATION SYSTEM

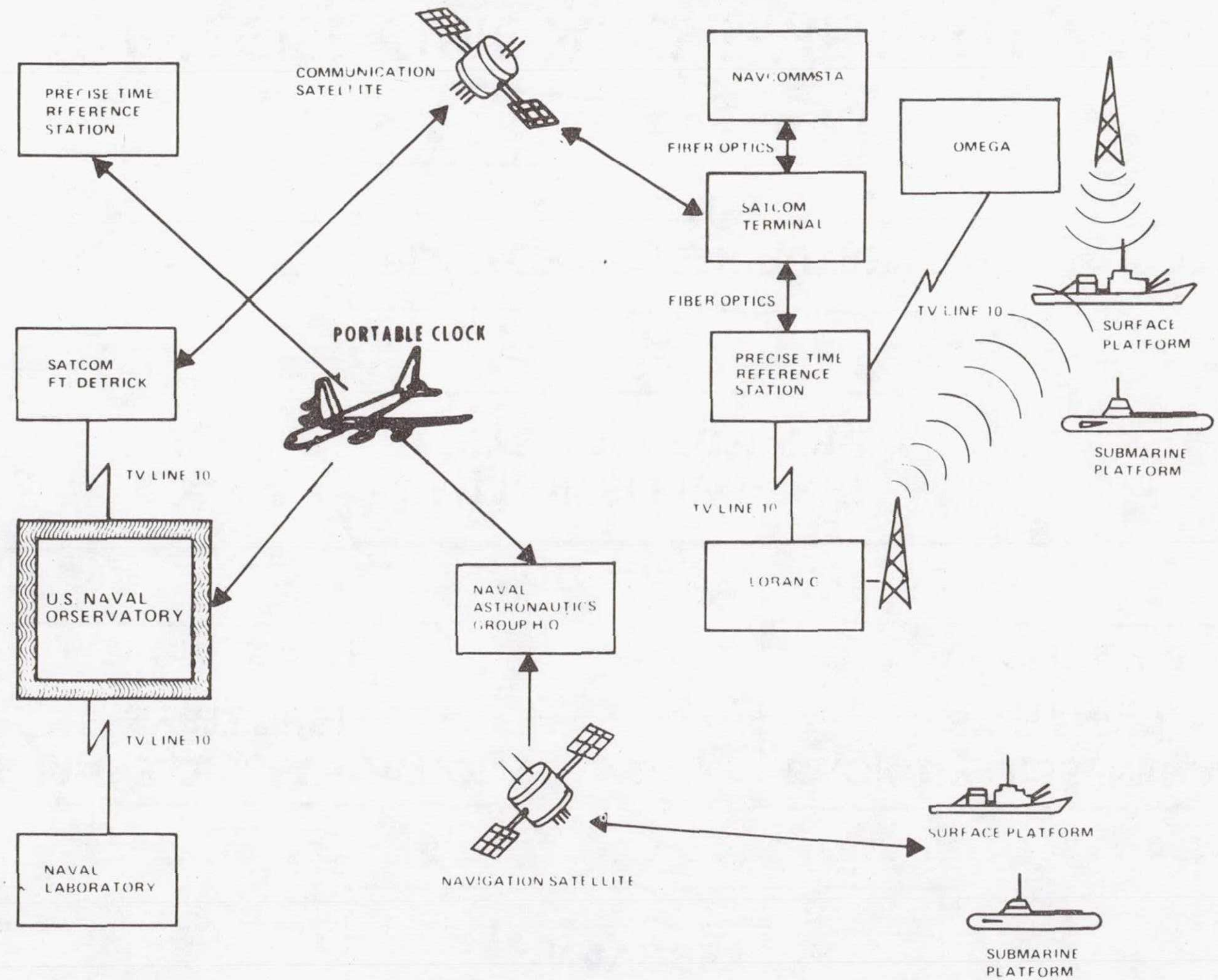


FIGURE 3

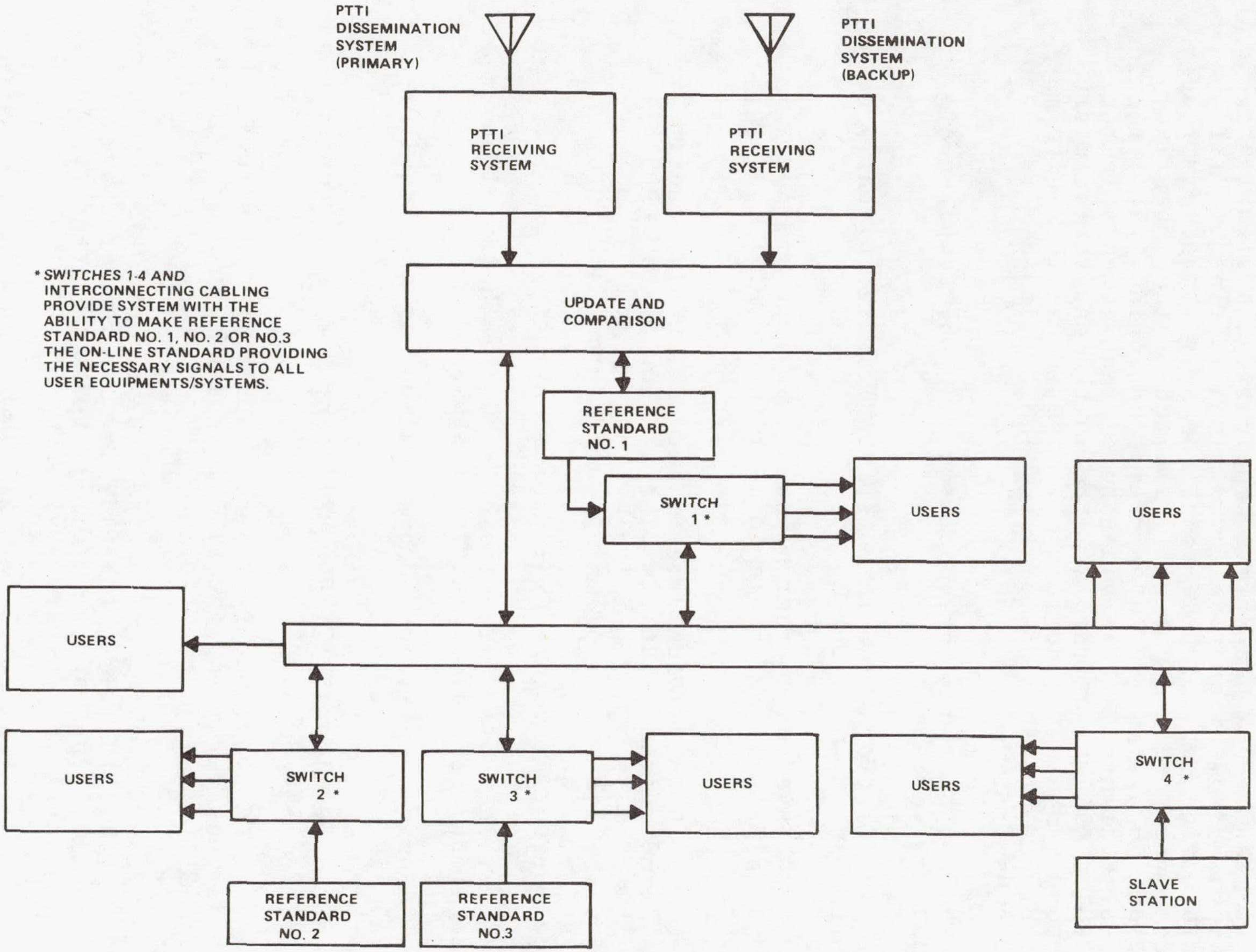


Figure 4 PTTI PLATFORM DISTRIBUTION SYSTEM CONCEPT

delivered to CNO for validation in November 1978. The Analysis indicates that there will be a dramatic increase in the number of platform communication, navigation and weapon systems requiring increasingly stringent PTTI information; e.g., while today the Navy's platform requirements for time are primarily limited to the submarine fleet and are in the 100 microsecond region, the typical Naval Platform, afloat and ashore, requirements in the 1980 to 1990 time frame will be in the one to ten microsecond range with one future system identified which may require a platform to maintain time to within $\pm$ ten nanoseconds or less.

Some of the major conclusions in the Requirements Analysis were that:

. Platform requirements for PTTI are currently being met.

. Based on current planning, surface platforms especially will be unable to meet future PTTI requirements.

. Those platforms requiring PTTI would benefit from a "Standardized" Platform Distribution System (PDS).

. With proper planning, all current and future platform requirements for PTTI can be satisfied by at least one current or future PTTI dissemination system.

. There is no identified need for the development of additional PTTI dissemination systems.

Some of the major recommendations in the Requirements Analysis were to:

. Develop a PTTI PDS utilizing a modular concept and standardized equipments.

. Consider the PTTI Requirements Analysis recommendations regarding the PTTI Dissemination System in any PTTI PDS development effort.

Based on the requirements analysis, CNO is developing a new PTTI OR. Also, the PTTI Program Master Plan is currently being updated and will be resubmitted to CNO for approval in February 1979.

Some of the major on-going efforts to be covered in

the updated Program Master Plan include the:

. PTTI Maintenance and Calibration Program.

. SSN-637, SSN-688 and POLARIS/POSEIDON Class PTTI PDS Developments.

. NAVOBSY Master Clock System Upgrade.

. AN/URQ-23 Frequency Standard.

. GPS Time Transfer Unit.

. PTTI Technology Program.

The PTTI Maintenance and Calibration Program at the Naval Electronic Systems Engineering Center (NESEC), Portsmouth is a continuing program. Responsibilities under this program include portable clock trips to various Navy and DSCS earth terminal sites and the repair and maintenance of all Navy rubidium and cesium beam frequency standards. Beginning in FY-79, this program will also have full responsibility, including budgeting, for the VERDIN 0-1695 Cesium Beam Frequency Standard Depot.

NAVELEX, in May 1977, developed a system level specification for a PTTI PDS for the SSN-637, SSN-688 and POLARIS/POSEIDON Class Submarines. This effort originated with an urgent need for PTTI on the SSN-637 Class Submarines. The specification calls for an automatic, no break system and for the use of common equipment in all of the classes with reconfiguration to meet the needs of each particular class.

NAVELEX is currently developing platform distribution systems for the SSN-637, SSN-688 and POSEIDON Class Submarines. The SSN-637 Class PTTI PDS depicted in Figure 5 is an automatic, no-break system. The PTTI PDS for the SSN-688 and POSEIDON classes, which is depicted in Figure 6 will essentially be in accordance with the original May 1977 specification, except that the need for an automatic, no break system has been eliminated by the program sponsor.

NAVELEX has also been tasked by CNO with upgrading the NAVOBSY Master Clock System, see Figure 7. As stated previously, the PTTI Requirements Analysis has indicated that current Navy requirements are in the 100 microsecond region and that future requirements will be in the one to

SIMPLIFIED DIAGRAM OF SSN-637 CLASS PTTI PDS

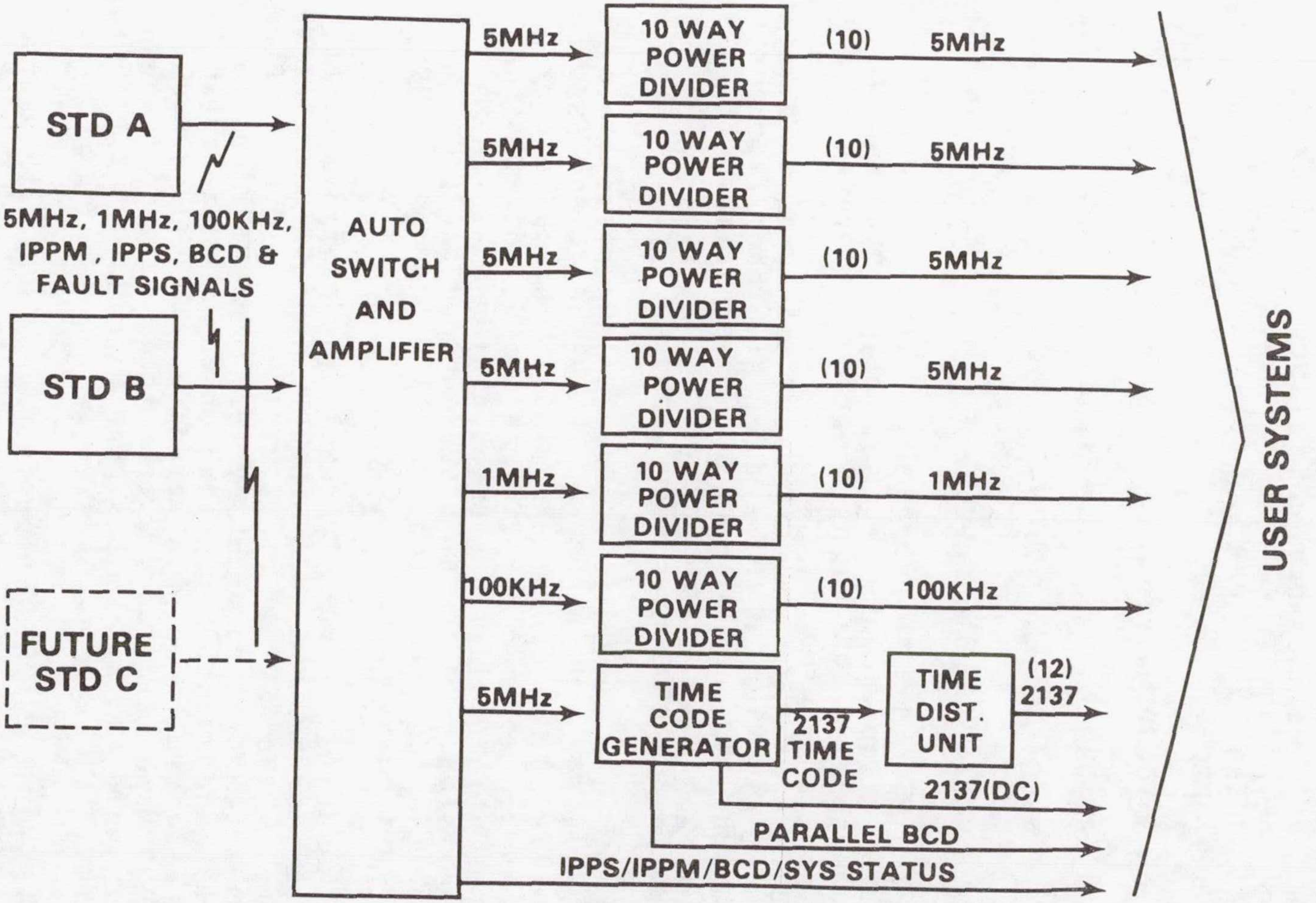


FIGURE 5

PCIP & SSN-688 PTTI PDS

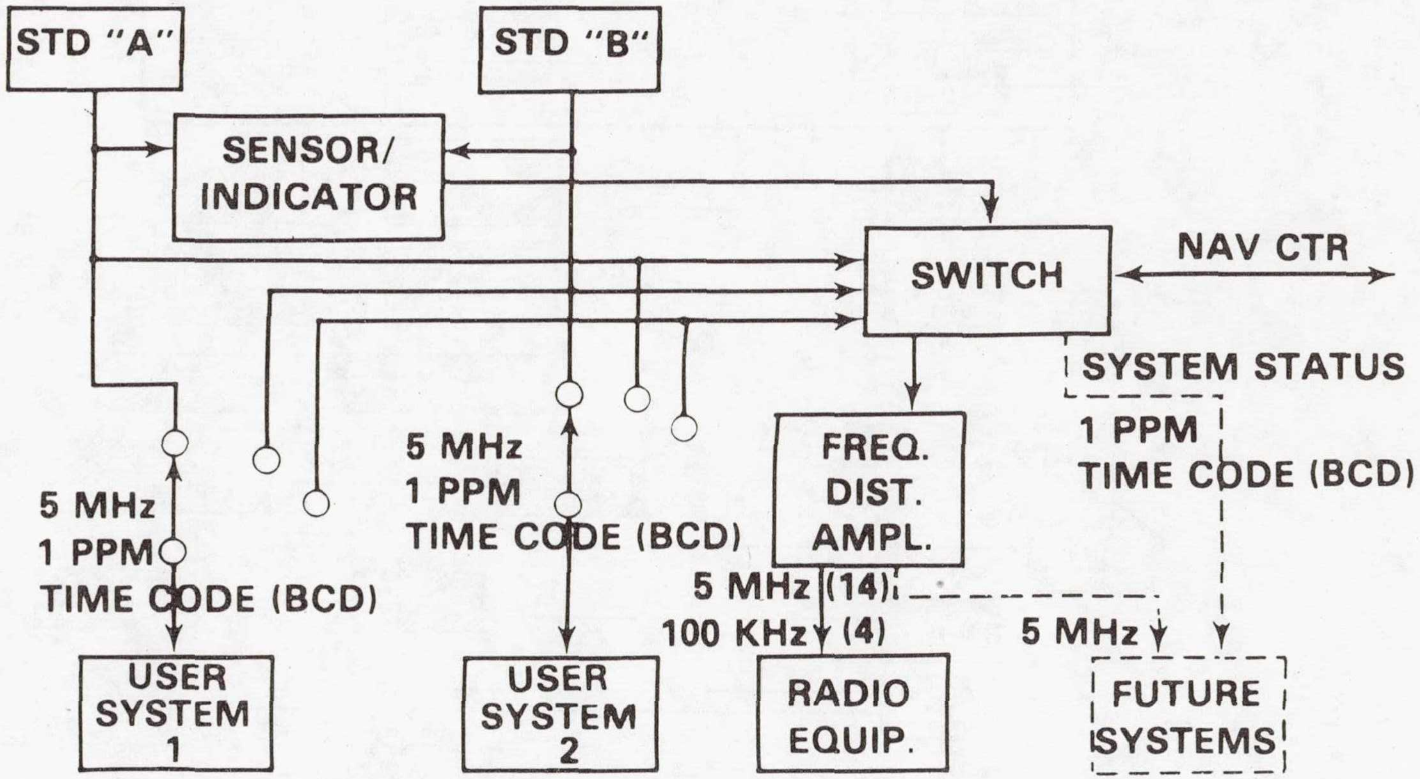


FIGURE 6

U.S. NAVAL OBSERVATORY MASTER CLOCK SYSTEM

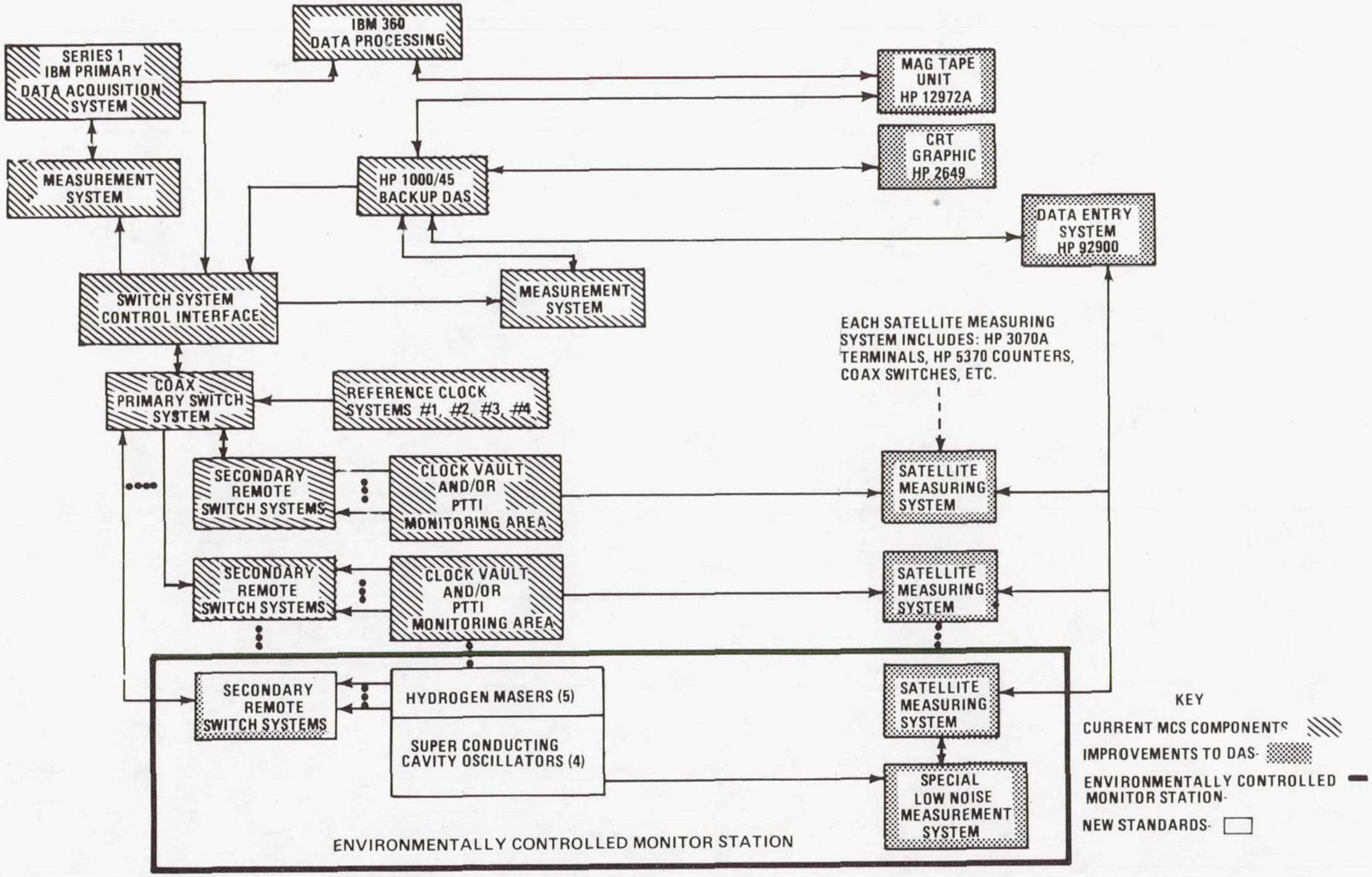


FIGURE 7

ten microsecond range with one requirement possibly in the ± ten nanosecond region.

A similar increase in PTTI requirements for other DOD components is also projected, e.g. while the NAVOBSY, under its assigned responsibilities as the single DOD component responsible for PTTI management control functions, currently maintains the Master Clock System which also serves as the DOD Master Clock and provides the National Security Agency, the Air Force Calibration Laboratories and the DSCS with time of day in the 0.1 to 10.0 microsecond range, it has projected requirements to provide time of day in support of the GPS and the Transit Improvement Programs to within 10 nanoseconds or less and to provide the Very Long Baseline Interferometry Program with a frequency stability of one part in 10^{14} or less.

While the present system provides time accurate to within 1 nanosecond per day and exhibits an exceedingly good long-term stability of two parts in 10^{14}, it will not meet the projected Navy and DOD user system requirements. In order to support those requirements, the future Master Clock System will be required to maintain time accurate to within 1 nanosecond per ten days and a frequency stability of 1 part in 10^{15} or less.

This will be accomplished by the introduction and integration of the latest, state-of-the-art technology and improved reference standards into the current Master Clock System to maintain and display near real time in the sub-nanosecond accuracy region.

This upgrade will enable Navy and other DOD components to meet their projected PTTI requirements and it will maintain the NAVOBSY's preeminent role in the national and international PTTI community.

While these basic documentation, program planning and early development efforts are taking place, there are still, however, real on-going operational needs of the fleet which must be met. For example, the AN/URQ-10 and 10A are the reference frequency standards of the fleet today. They are approximately ten years old. Although they still meet the majority of the Navy's current PTTI requirements, spare parts are becoming increasingly expensive and unavailable and handling procedures for maintenance are out of date.

The AN/URQ-23, disciplined time and frequency standard, which will be a direct replacement for the AN/URQ-10 and 10A was granted Provisional Approval for Service Use in September 1977. It offers an improved frequency stability and timing information which the AN/URQ-10's do not. More significantly, however, is the fact that the AN/URQ-23 can be calibrated in four hours vice the approximately 30 to 60 days required for the AN/URQ-10's. This calibration feature alone should yield significant savings for the fleet.

The AN/URQ-23 is currently undergoing First Article Testing at NRL and full Approval for Service Use is expected in June 1979.

Also, at the direction of CNO, NAVELEX is procuring a GPS Time Transfer Unit (TTU) Feasibility Model, see Figure 8, for the NAVOBSY via the Space and Missile Systems Organization (SAMSO).

The GPS has an inherent ability to disseminate PTTI information. The Time Transfer Unit, on which a paper was given at the Ninth Annual PTTI Applications and Planning Meeting (1), should provide fixed platforms with a low cost means of extracting that information and updating their local clocks to within 100 NS of GPS time which in turn is referenced to the NAVOBSY.

Finally, NAVELEX was recently notified that the Naval Material Command (NAVMAT) is partially funding the PTTI technology effort outlined in Figure 9. Informal discussions with NAVMAT and the NAVOBSY indicate that the PTTI technology effort has enough potential impact on, for example, future atomic clock developments and PTTI dissemination via satellite clocks to warrant its continuation.

NAVELEX is, therefore, currently developing a PTTI Technology Program, with particular emphasis on (1) relativity effects on satellite clocks, (2) laser reflectometry for increased time dissemination accuracy via GPS and (3) super crystal technology, in support of both current and projected Navy PTTI Program requirements.

That completes the review of the most significant, on-going PTTI efforts. Some of the projected PTTI efforts to be addressed in the updated PTTI Program Master Plan include the:

GLOBAL POSITIONING SYSTEM
TIME TRANSFER UNIT

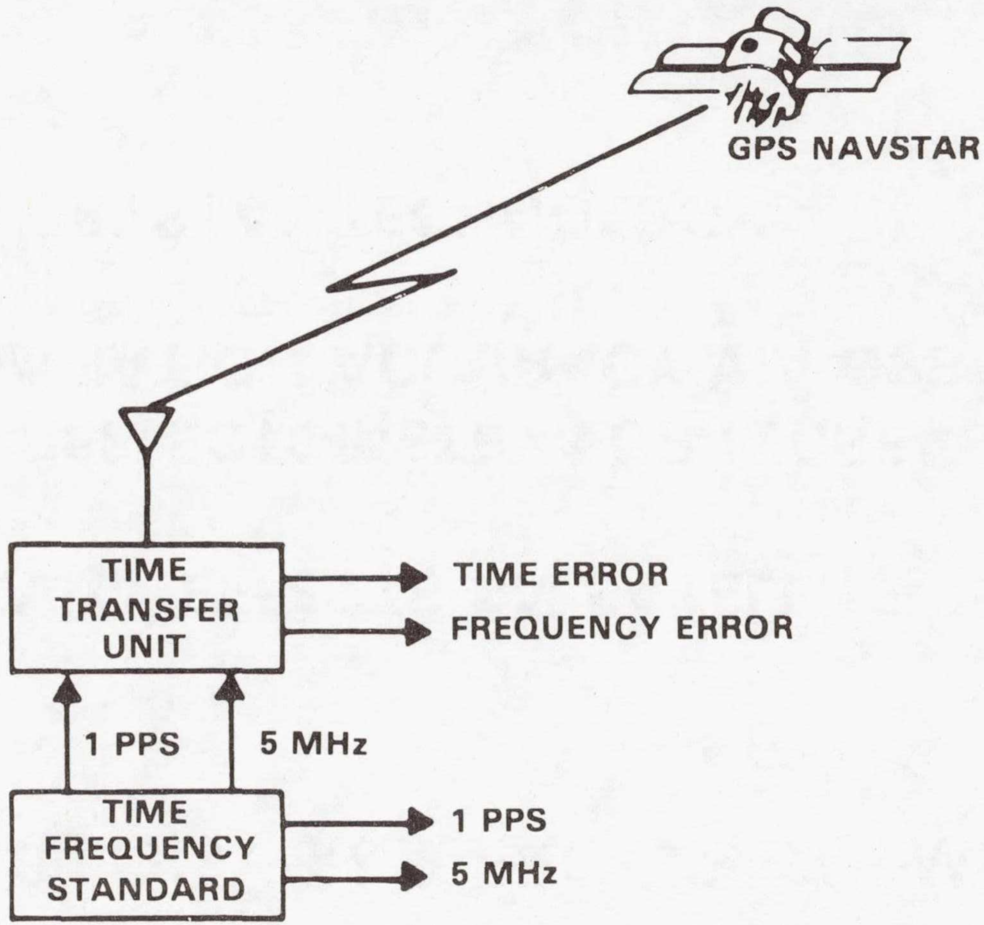


FIGURE 8

MAJOR PTTI
TECHNOLOGY R&D PROGRAM EMPHASIS

- **RELATIVITY EFFECTS ON SATELLITE CLOCKS**
- **LASER REFLECTOMETRY FOR INCREASED TIME DISSEMINATION ACCURACY VIA GPS**
- **SUPER CRYSTAL TECHNOLOGY**

FIGURE 9

. 0-1695 Cesium Beam Standard

. Rubidium Standard Development

. "Standardized" PTTI PDS Development

. "Standardized" PTTI PDS Preliminary Design and Cost Analysis

. PTTI via DSCS

Under the VERDIN Program, NAVELEX several years ago developed the 0-1695 Cesium Beam Frequency Standard and a complete support system. Within the last year, however, agreement has been reached that cognizance of the 0-1695 and its previously mentioned depot responsibilities properly belong under the Navy's PTTI Program. NAVELEX is currently planning for that transition in FY-79.

The Naval Air Systems Command (NAVAIR) is currently preparing to develop a rubidium frequency standard in support of the TACAMO Program. NAVELEX under a previous program was funding for the development of a rubidium standard. NAVAIR has been contacted to see if the NAVELEX rubidium standard effort will meet the TACAMO requirements. If so, it is felt that substantial savings in development costs can be accrued by the Navy.

The "Standardized" PTTI Platform Distribution System Development, depicted in Figure 10, is a FY-80 new start effort to provide all Navy platforms, surface, subsurface, air and shore, which require PTTI information with a modular distribution system comprised of "Standardized" equipments which can be reconfigured to meet the needs of a particular platform.

The "Standardized" PTTI Platform Distribution System Development will take advantage of any outputs from the previously mentioned submarine efforts by incorporating them, to the maximum extent possible, into the "Standardized" system.

The first task under the "Standardized" PDS effort is a preliminary design and cost analysis to provide answers to some of the questions raised as a result of the PTTI Requirements Analysis; e.g.

a. How much standardization of equipment components

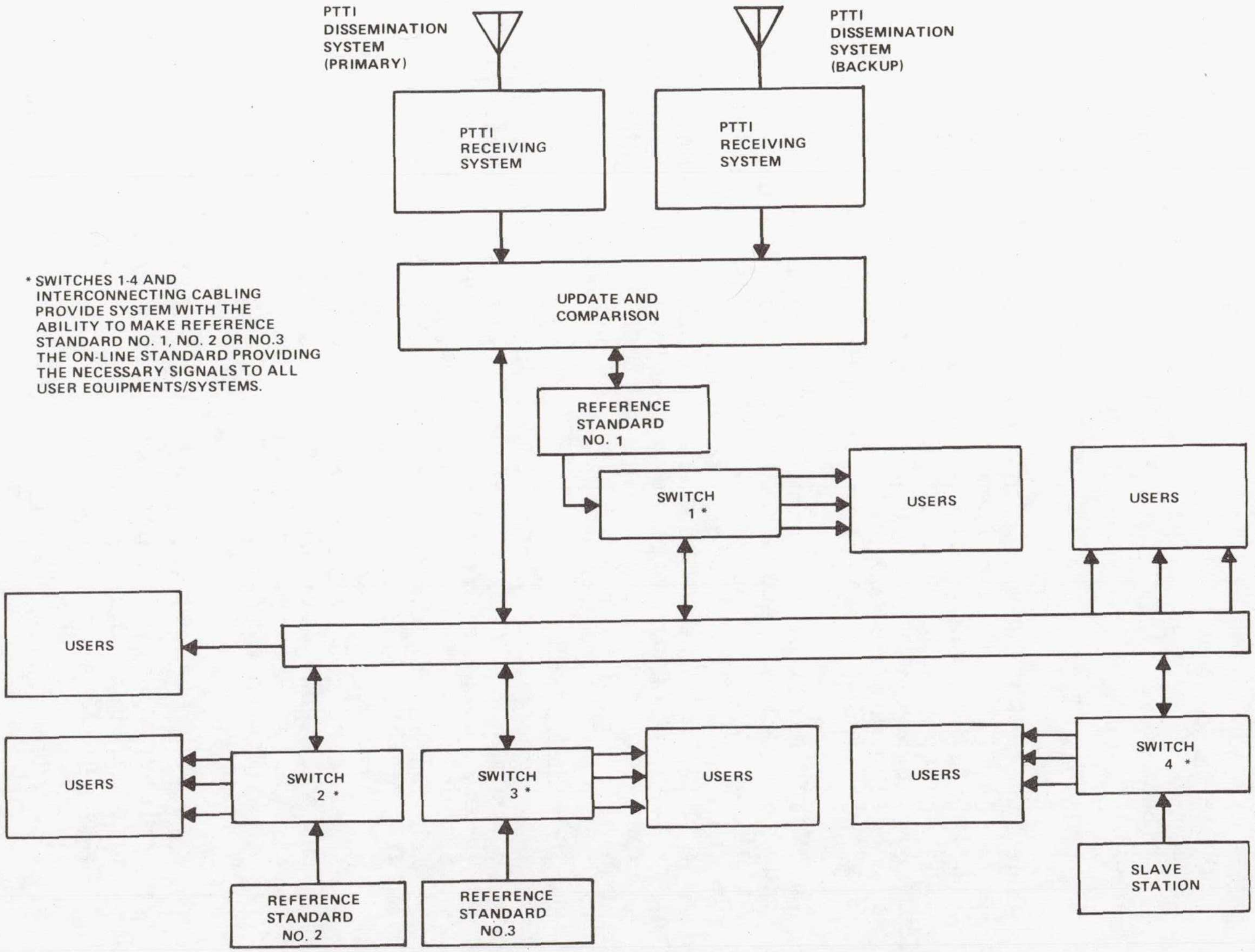


Figure 10 PTTI PLATFORM DISTRIBUTION SUSTEM CONCEPT

can be achieved between surface, subsurface, shore and air platforms?

b. How much utilization of existing assets is practical?

c. How much savings, in the area of logistics support, could be accrued by the Navy from the standardization of PTTI component equipments?

The final projected effort to be discussed is the PTTI via DSCS effort, see Figure 11.

The Joint Chiefs of Staff (JCS) on 17 March 1976 authorized the dissemination of PTTI via all DSCS terminals equipped with specialized modems capable of transferring time. The JCS further requested that CNO develop an implementation plan for the dissemination program.

The implementation plan was developed by the Naval Telecommunications Command (NAVTELCOM) in conjunction with the NAVOBSY and NAVELEX in May 1977. It designates the CNM as the executive agent for procurement and logistics support for this tri-service effort. CNO approved and forwarded this implementation plan to JCS for approval in September 1977. NAVELEX is currently planning for the assumption of these responsibilities upon the approval of the Implementation Plan by JCS.

As the preceding discourse indicates, the Navy's PTTI Program has, in three years, been defined and given direction. As Figure 12 indicates, however, the Program's major development and procurement efforts are still to come.

While there are problems associated with some of these efforts, e.g. the AN/URQ-10 maintenance problems, solutions have generally been identified and are being pursued.

The estimated impact of the outyear hardware procurements resulting from the platform distribution system development efforts and the JCS approval of the DSCS Implementation Plan will be significant. NAVELEX, CNO OPS 952, the PTTI sponsor, and 986, and COMNAVTELCOM are working together to absorb that impact.

The overall picture for PTTI has already improved significantly in the past three years; e.g. with the

TIME TRANSFER BY COMMUNICATION SATELLITE

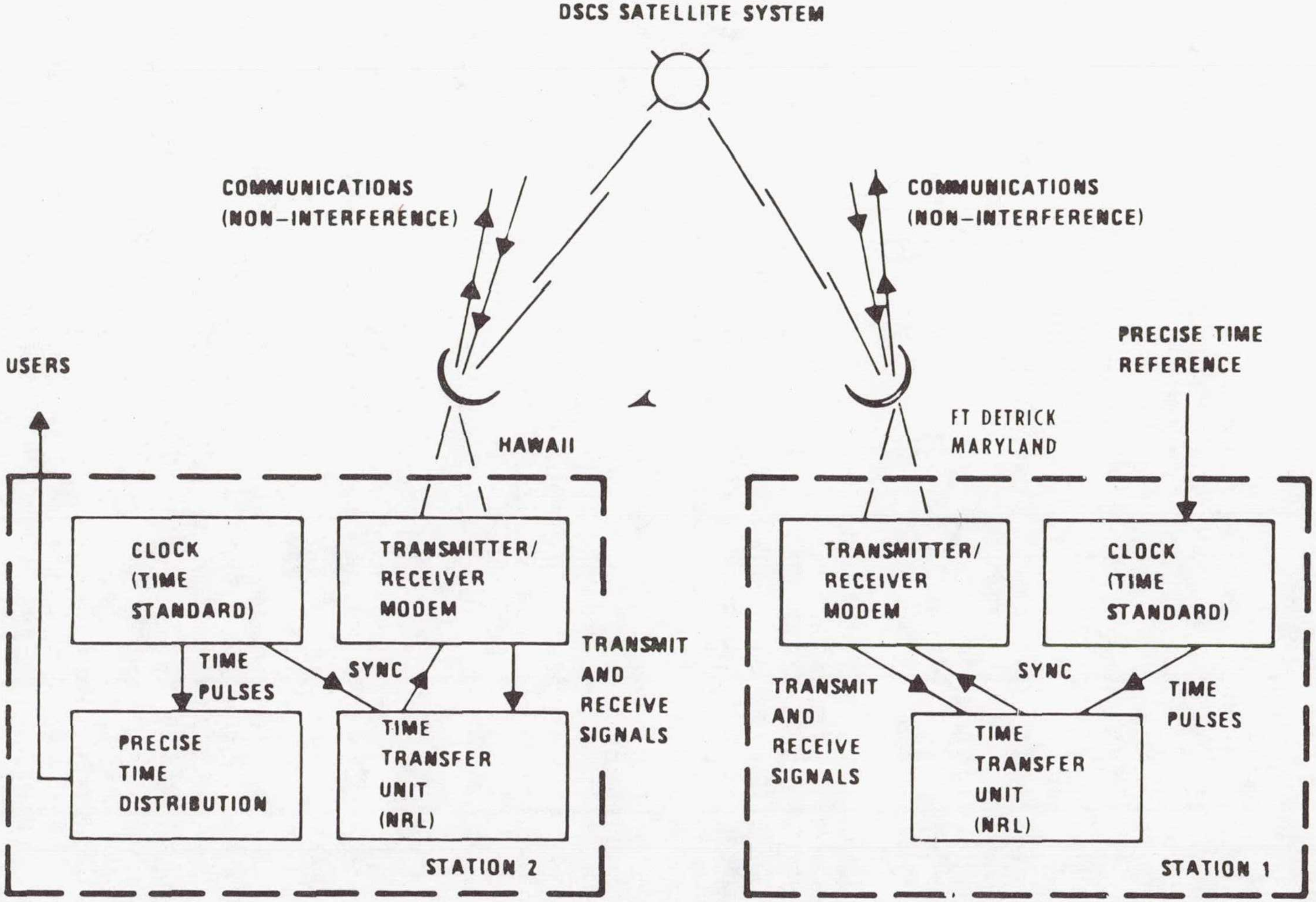


FIGURE 11

MAJOR PTTI PROGRAM MILESTONES

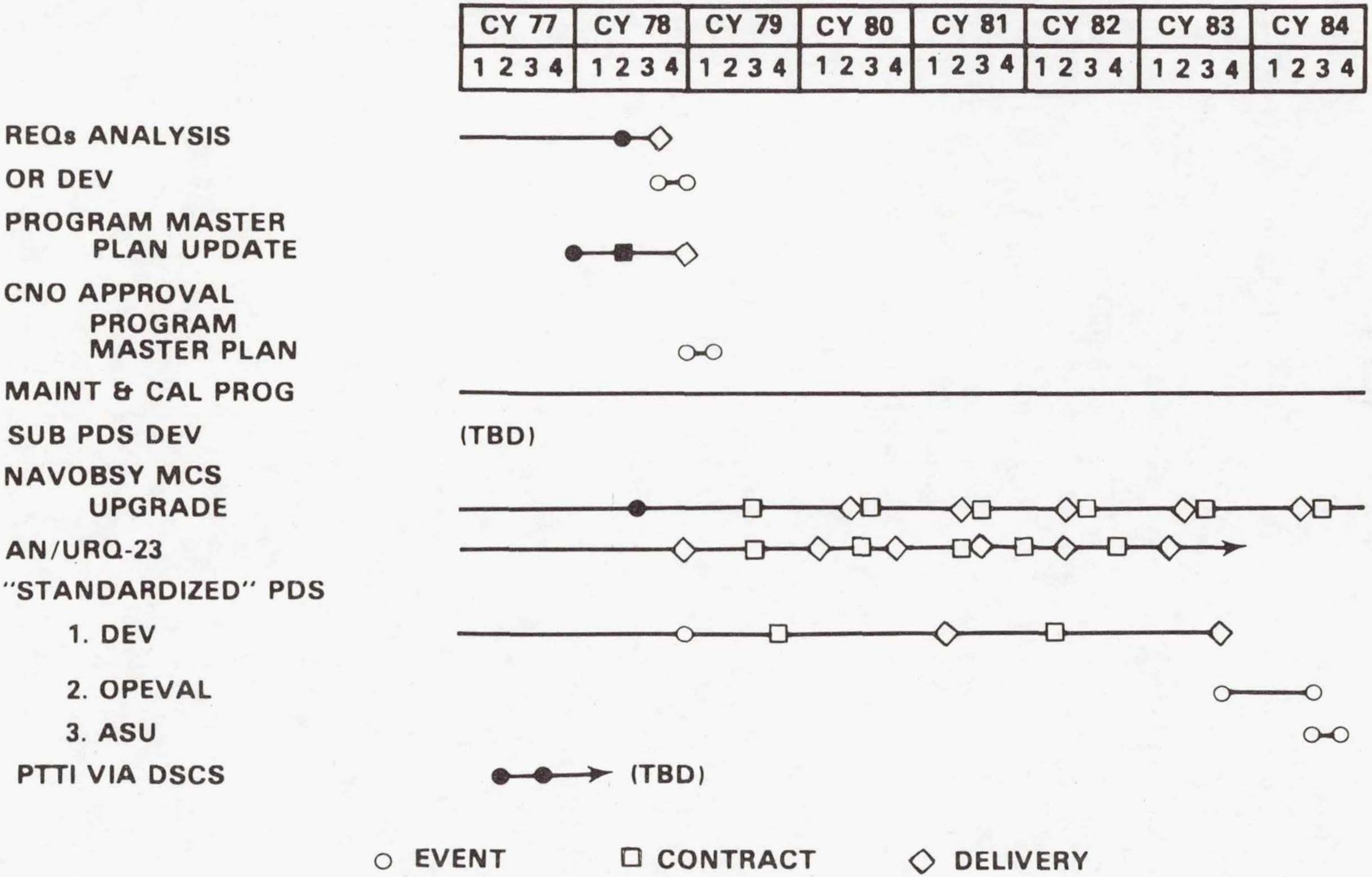


FIGURE 12

advent of the PTTI Program Master Plan, the PTTI Requirements Analysis and the assistance of OPS 952 and 986, the "Standardized" PTTI Platform Distribution System Development has moved from a new start in FY-82 to FY-80.

Again, in summary, the Navy's PTTI Program has, in three years, been defined and given direction. Program schedules and funding have improved and solutions have been identified for problem areas such as the AN/URQ-10. Overall the future of the program looks good. It is now being recognized by Navy Headquarters and Field Activities, the Fleet and other DOD Activities as attempting to solve the PTTI problems of today and to plan for the PTTI needs of tomorrow.

REFERENCES

1. Witherspoon, Jackson T. and Schuchman, Leonard, "A Time Transfer Unit for GPS," Proceedings of the Ninth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, March 19, 1978.

QUESTIONS AND ANSWERS

MR. HARRY PETERS, Sigma Tau Corporation:

I don't know how appropriate this is, but a couple of years ago there were ongoing efforts by Hughes Aircraft Company, RCA, Naval Research Labs, Smithsonian Astrophysics Lab, to get PTTI using hydrogen masers. Now we haven't heard much about that lately, nor do I see very much here about the future of this effort. I wonder if you could elaborate on that?

MR. ALLEN:

I believe those efforts are under the GPS program. And they are not directly under the PTTI program. We are trying to absorb more and more of the PTTI efforts in the Navy and bring them under one single program. But, again, this is not a GPS effort right now. Someone from the GPS program office would have to address those development efforts.

MR. CLARK WARDRIP, NASA Goddard Space Flight Center:

When do you see these GPS timing receivers becoming available?

MR. ALLEN:

The feasibility model is supposed to be delivered, I believe, next June to the Naval Observatory. And I don't see a unit being available much before a year after that.

MR. WARDRIP: Can you give an estimate of the cost?

MR. ALLEN:

No. Estimates have ranged anywhere from about $25,000 to $50,000, but I think until we have the feasibility model and do some testing, we really won't be able to put a price on it.

MR. DAVID ALLAN, National Bureau of Standards:

It is interesting that at the one nanosecond level on a global basis the relativistic effects become extremely important. And to date, internationally, a particular coordinate system that would be useful for this has not been agreed upon. And so this is a separate issue outside of this audience that needs to be flagged and watched. I am glad to see that you are very conscious of that in your considerations.

DR. GERNOT M. R. WINKLER, U. S. Naval Observatory:

That is exactly why you have seen in this program a combination of relativity concerns.

MR. ALLEN: Yes.

FREQUENCY DOMAIN MEASUREMENT SYSTEMS

Michael C. Fischer
Hewlett-Packard Co.
Santa Clara, Ca.

ABSTRACT

Stable frequency sources and signal processing blocks must often be characterized by their noise spectra, both discrete and random, in the frequency domain. In this tutorial, conventional measures are outlined, and systems for performing the measurements are described. Broad coverage of system configurations which have been found useful is given. Their functioning and areas of application are discussed briefly. Particular attention is given to some of the potential error sources in the measurement procedures, system configurations, double-balanced-mixer-phase-detectors and application of measuring instruments. A general calibration scheme is detailed.

INTRODUCTION

This paper is a lightly edited transcription of the tutorial presentation which included audience participation. With that excuse, the author begs the reader's indulgence of the conversational style and disorder.

First, to point out the areas that make up the subject of frequency domain measurement systems, see Table 1. This is only roughly the order in which they will be covered. References will be cited for coverage of the basics.

Model numbers, ranges and accuracies of the various frequency domain instruments will be avoided since these are readily available on data sheets and in catalogs. However, there are some traps in the ways these instruments may be used which will be covered.

Measures

To establish a clear context, the symbols and definitions that are used for the measures are shown in Table 2. The reader may be familiar with one or more of these from various applications. Basically, each of these definitions implies a math model of a random process. When a measurement is made, it will most likely include processes which are both random and non-random - an important source of some of the traps. Accordingly, it is necessary to have a clear understanding of how a chosen measure (math model) responds to discrete spurious components as well as continuous random spectra, and as separate questions, how the measurement system relates to the measure for both classes of signals. These areas are covered in references 1, 2 and 3.

Systems

I would like to pose a question to start into the systems aspects of the subject. Why do we need a system to measure frequency domain stability? (see Figure 1). The usual reason is dynamic range. We might try to use a spectrum analyzer to measure the noise side bands, for example, which are 140 dB below the carrier in a 1 Hz bandwidth. Even if we use a 1 kHz bandwidth in the spectrum analyzer instrument, this noise will be 110 dB below the carrier as indicated on the screen of the analyzer, and the skirts of the one kHz filter would force us to look no closer than several kHz away from the carrier. So the dynamic range in both the frequency and amplitude sense, and their interaction, are the reasons why we have to go to more complexity than this, with higher quality oscillators. There is a large family of oscillators for which this is an appropriate way of measuring their noise side bands, but it displays AM and PM.

AM versus PM is another distinction which must be cared for. The usual name for the measure that is seen on a spectrum analyzer is the rf power spectrum. This measure is almost never used as a specification for a frequency standard or any component of a frequency distribution system. What is needed is a demodulator, a phase or frequency demodulator, implying that it is relatively insensitive to AM and it needs to be highly sensitive to small PM for FM. (See Figure 2). Otherwise, in the case of those rare systems that have AM sensitivity, it should probably be specified separately since there would be a great opportunity for confusion.

Referring to the instability sensor in the middle of Figure 2, the demodulator should have a very good ratio between the demodulated signal and its own noise contribution. We could apply some other words to that central block: an error enhancer, an error multiplier (which frequency multipliers tend to function as) or conversion to base band. In many communication systems, some element of, or some of the blocks of the system itself (which is going to use the oscillator under test) might be applied here. It is important that their transfer function be well understood and modeled. If so, they may be applied as the instability sensor to make the measurement in a very realistic fashion with respect to the oscillator's performance in the final system.

Autocorrelator System

The reason for choosing the fairly complex looking system of Figure 3 as the first to be considered in detail is that it has wide application. In many of its forms, it can take a range of input frequencies without modifying the hardware, and it can test a single source without having a reference source,without requiring a second similar source as a reference. This feature makes it attractive for development work in many programs. Figure 3 may be explained by starting at its output and working backwards. The double balanced mixer is used as a phase detector. The fundamental operation that the mixer performs can be described by assuming sine waves applied to the R and L ports. The simplest model for the mixer is that it multiplies the two sinusoids together so that what comes out are two signals as in the trigonometric identity. One signal is an average term which is proportional to the phase difference between the two input sine waves (that is the difference frequency out of the mixer). The other is a very large signal which is the sum frequency output of the mixer. Since the inputs are the same frequency, this sum output is twice the input frequency. The phase shifter, shown here in what may be called the reference channel, is adjusted to bring the two mixer input signals into quadature so that the mixer performs a phase detection. If, on the other hand, the two input signals are brought into in-phase condition (or 180^{o} of out phase) by the use of this phase shifter, then this becomes an amplitude detector, and the AM spectrum of the input signal can be analyzed.

The functions in the other path to the mixer cause this to

be an instability enhancer or a detector of the spectrum that we want to measure. First, the directional coupler is shown as having its 10 dB attenuating port to the L port because we expect to take less loss down this path and we want as large a signal as possible to both ports of the mixer. The result with any one of the items listed on the right of Figure 3 is that this path functions as frequency-to-phase transducer. For example, in the case of the transmission line, the signal goes down the transmission line, reflects and comes back, incurring a delay. The delay of 300 feet of coaxial cable at 1 MHz is approximately one cycle; 1/2 cycle out, and 1/2 back. If the frequency then changed by 10%, the phase shift through this path length would change by 10% of a cycle. In that way, a delay line or a narrower filter on this path serves as a frequency-to-phase transducer. Then the mixer acts as a phase-to-voltage transducer. The overall result is a sensor which transforms frequency to voltage, a low noise discriminator.

There is an interesting aspect of the narrower band resonators in this system - either a crystal filter or a cavity. Within the bandwidth of the resonator, all the foregoing is still true. However, outside the bandwidth of the resonator, the resonator can be considered to be stripping off the noise sidebands, or holding the phase of the signal through this path coherent for a longer time than the period of these larger sideband or modulation frequencies. If a phase change occurs at a high rate versus a narrow resonator, the phase out of the resonator does not change, it does not follow the rapid phase change, and the system output is now the phase modulation spectrum, outside the bandwidth of the resonator.

For areas of applicability, the crystal filter, from 100 kHz to the 100 MHz region is useful. The one port cavity is best applicable from 5 GHz up. One could use a transmission cavity similarly. The long, shorted transmission line is the broadband approach which allows the single system to operate from 100 MHz to 10 Ghz. This entire system does not offer as low a system noise floor as some of the other systems, but it does have extremely broad applicability and the advantage of single input source. The references appropriate to Figure 3 are: 4, 5 and 6, the last of which dealt with this system using a crystal as a filtering element.

Preamplifiers

The modulation spectrum appearing at the mixer output in

these systems still must be measured by a frequency domain instrument such as a low frequency spectrum analyzer. The sensitivity and noise figure of these instruments usually requires a preamplifier following the mixer to give best performance. The mixer will generally be a 50 ohm component which means it was designed to operate from 50 ohm sources into a 50 load, not that it looks like 50 ohms in general, particularly when both ports are driven from a relatively high level. The low noise amplifier, which is needed to achieve good performance from the system, is trying to operate from a source impedance which is probably below 50 ohms. The optimum-noise source resistance for low frequency amplifiers is very rarely this low. The only one that is commercially available which the author has seen referenced (but has no personal experience) had the brand name Ortec mentioned in Lances paper at 77 PTTI. For that purpose, the author has built his own relatively simple amplifiers. A good reference on building amplifiers of this class (and low noise rf amplifiers also) is a book by Motchenbacher and Fitchen, reference: 7 on low noise techniques.

Mixer-Phase Detectors

Since the mixer is a critical element, let us digress from the systems configuration for a moment and give the mixer more attention. The author has seen many systems where people have attenuators preceeding mixer, even following the mixer, and expecting that when they click in 20 dB, to have a 20 dB change in what is going on in the system. This is probably not true because of the way the mixer works. In Figure 4 the diagram on the left is the schematic of a good many mixers. In fact, some of the best mixers for our purposes have exactly that schematic. There are also many variations, for which this is still a good first-order model. Even though their schematics look vastly more complex than this, they look, to the input and output signals, very much as this model.

Though the results are usually misleading, it is natural to attempt to measure the drive level to the mixer by looking at the voltage at its input ports. It is very easy to put a Tee connector on the BNC going into the mixer and probe the signal at that point with a good high impedance scope probe as was done for Figure 5. But, just as one wouldn't try to determine the dissipation in a zener by measuring the voltage across it and ignoring the current, one should not draw many conclusions from this voltage. A high impedance rf voltmeter looking at the voltage into

the mixer is even less informative about what is going on in the mixer because all that is being measured is the forward drop across a diode. As is indicated by the functional model in the center of Figure 4, which assumes that the L port signal is the larger one, it simply saturates the mixer. With a large L signal, one series pair of diodes or the other, for either polarity signal, is strongly forward biased. This amounts then to connecting either one side, one polarity, or the other, of the R signal to the X port determined by the polarity of the L signal. The results of this action are shown in the right-hand part of Figure 4 for R and L in a quadrature phase relationship. Corresponding to this function, the double balanced mixer is sometimes also called a reversing switch. It is interesting that there is very little literature on the mixer used as phase detector. References:8 and 9.

Figure 5 shows mixer waveforms which are careful tracings of scope photos. Both the L and R port waveforms appeared to be identical. Both were driven with + 7 dBm incident power, in quadrature phase. That was measured with the coax disconnected from the mixer, into a good 50 ohm load, with an rf power meter. Then the coax was reconnected onto the mixer and probed with a high impedance probe for the photo. Notice that the waveform is distorted by the action of the diodes. Another important point is that this signal came from a broad band resistive 50 Ω source. This is probably not true of most of the output stages of the frequency standard devices which might be connected to a mixer. Now, given that the signal came from a 50 ohm source, and was sinusoidal into a 50 Ω load, the mixer looks like various things. It may look like more than 50 Ω as the input is crossing zero, and it clearly looks like less than 50 Ω as the diodes start to conduct more and more. The mixer, at this drive level, probably drops down to the order of 20 Ω or below, over portions of the waveform.

At the output X port the 2f output or sum frequency appears. The important thing to notice is how large it is relative to the inputs. In other words, when a mixer is used in one of these systems, one must be aware of the fact that with a 5 MHz input signal to the mixer, there will be a very large 10 MHz output signal from the mixer. That large output signal can deal a very strange blow to the spectrum measuring instrument, or to the low noise amplifier that is connected to the mixer. Although the X output does not look exactly sinusoidal, spectral analysis of that waveform at the bottom of Figure 5, shows that its third harmonic is still 30 db below the fundamental. The top of the

screen was set to equal the power incident on the mixer at 10 MHz, so the output signal is about 6 db below the power that is incident on the inputs.

Mixer Loads and Sensitivities

Most systems will require a low pass filter following the mixer to reject the large 2f output signal. Most familiar filter designs have input impedances which are a large mismatch outside their passband. That is to say that they achieve attenuation by reflecting the unwanted input. A low loss filter, made of only reactive elements, must then appear as a highly reactive mismatch at its input to signal frequencies outside the passband. This situation could result in the mixer being terminated reactively at the frequency of its largest output signal component. This condition has been seen to cause surprisingly gross distortions of the input waveforms. This can raise serious questions of the validity of the measurement because the device under test is driving a mixer input which is an extremely non-typical load, having a large reactive mismatch and violent non-linearities.

The low pass filter following the mixer can be designed to present a 50 Ω load to the 2f output component by making the input element of the filter a shunt capacitance with a 50 Ω resistance in series. Figure 6 is an example from reference 10. This arrangement terminates the mixer in a non-reactive 50 Ω for all frequencies above 1 MHz, but unloads the mixer to maximize its sensitivity below 100 kHz. Since the filter element values depend on seeing a 50 Ω source resistance in the mixer, this filter develops one to two decibels of peaking when the mixer is driven above 0 dBm at both inputs. This is a good example showing the necessity for checking system flatness, over the entire frequency range to be measured, during the calibration procedure.

The transducer coefficient of the mixer, used as a phase detector, in volts per radian, (that is, its ratio of converting radians of phase shift to volts of output signal, dc average, or low frequency average) can be quite sensitive to the terminating impedance at its output port. See Figure 7. Taking the output into a broadband 50 Ω load as a base line, the output slope can be increased up to 6 dB by raising the load resistance. However, another 6 dB, for a total of 12 dB, approximately, increase in sensitivity can be gained by careful choice of a parallel capacitive reactance well below 10 Ω (at the input

frequency) as a termination, reference: 11. Again, being a reactive termination, this can raise the above questions due to mixer input waveform distortion, and has been shown to roll off the modulation frequency response, requiring thorough calibration.

All but one of the curves in Figure 7 were taken with the same RF power level applied to both the L and R ports. Using an HP 3335A, a calibrated phase shift, was inserted and the static transducer coefficient was measured using a DVM. One-tenth radian positive and one-tenth radian negative about 0 were the phase shifts used. Note that if the mixer is terminated at a high impedance, a gain from 4 dB, to much more than 6 dB is realized at lower levels, versus terminating the mixer with a broadband 50 Ω. And there is an interesting fact here that all the curves for all the mixers, at least at some input power levels, were asymptotic to the line whose equation is K_ϕ = 1.8 times the incident power level expressed in rms voltage; and the relationship seems to be that unloading doubles the peak voltage available out of the mixer. The lower frequency average is $2\sqrt{2}\ (2/\pi)$ which is 1.8.

This completes the digression to mixer details and attention returns to other ways of configuring the measurement system.

More Measurement Systems

Figure 8 shows a two-channel version of Figure 3, somewhat simplified to fit the page. This takes two of the systems of Figure 3 and splits the power of the oscillator under test into them. This allows the output signals to be cross-correlated for noise reduction. This is now possible because there are very convenient fast Fourier transform instruments which have two inputs. The Fourier transform has both amplitude and phase and it requires two inputs to the processor in order to accomplish that. There is a button on the front panel which commands a cross correlation between those two inputs. This, with averaging, will allow reduction of the noise of amplifiers which might be inserted. The noise of one amplifier is uncorrelated versus the noise of the other amplifier, allowing this improvement in system noise figure by using one of the more recently available fast Fourier instruments as the frequency domain analyzer for the spectrum. Reference: 12. Turning now from the single oscillator to the two oscillator systems, Figure 9 shows the most common system that is used for

frequency domain measurements. It requires a pair of oscillators of similar quality, unless the reference oscillator is much better than the one to be measured. This system has been discussed fairly extensively in the literature; references: 11, 13, 14. This, as far as I am aware, has been in wide use since 1964 for specifically this purpose. As the notations in Figure 9 indicate, the modulation spectrum coming out of the mixer is a phase modulation, outside the bandwidth of the lock loop, and a frequency modulation spectrum, inside the bandwidth of the lock loop.

This system can be modified slightly, as shown in Figure 10, to make measurements of a two port device, such as an amplifier. If a synthesizer or frequency multiplier is to be tested, which changes the frequency from its input to its output, then a similar device would have to be placed in both paths so that the same frequency goes to both inputs of the mixer. The phase shifter is used to bring the phases into quadature at the mixer. Most of the noise from the reference oscillator cancels since it appears at both inputs to the mixer. Again, reference: 11 is suggested.

The cross correlation enhancement of the measurement system can be applied to this system also, as shown in Figure 11. Here the system of Figure 9 has been duplicated, amplifiers can be inserted, and their noise can be suppressed by averaging in the cross correlation process. This was suggested with an analog multiplier as a correlator in reference 15, along with other systems considerations which are important to this kind of measurement. The dashed line shows that one of the loops can be used to lock the reference oscillator to the oscillator under test, to maintain quadrature phase at the mixers.

Up to this point the systems that have been considered either measured a single source or the combined noise of two sources (one considered to be a known reference) at the same frequency. In the systems to follow, the use of a reference source whose frequency is offset from the unit under test allows the use of period or frequency counting as the measuring instrument.

Figure 12 shows the simplest of these systems. In addition to the convenience of using a commonly available instrument, a period counter, this system has the advantage of allowing measurement of time domain stability, $\sigma_y(\tau)$, as well as aging or other drifts. Further advantages are that the

measurement resolution (dynamic range) can be extreme (even with an inexpensive counter) and that the measured raw data is in digital form, being stable and convenient for automation via an interface bus. References: 16, 17.

Since counter input stages usually do not have low enough noise to avoid degrading the signal level available from the mixer, preamplification is necessary. Optimal characteristics for this preamp are that it should have high gain in order to hard limit on millivolt inputs, low enough noise to add no more than sub-microsecond perturbations to the zero crossing of a 1.0 Hz, 1.0 Vp-p sine wave input, and have 1.0 MHz or better bandwidth. This is already a very specialized design, and further needs to have a calibrated (even adjustable) first stage bandwidth.

Recall that the output from the mixer will have a large component at twice the input frequency riding on the low frequency waveform whose period is to be measured. This requires a low pass filter of typically greater than 100 dB rejection, at twice the lowest mixer input frequency, because these perturbations are essentially uncorrelated with the unperturbed time of the zero-crossing and therefore function as a noise source. This can typically be accomplished with as few as four poles. However, the physical construction of the mixer-filter-amplifier combination has to be very sophisticated in order to obtain the attenuation that the filter was designed to provide. One of the problems in this field is that there are few systems for sale which are specified to do the overall measurement job.

An extension of the above technique is shown in Figure 13. This dual mixer time difference system also happens to take data in the time domain, as do most Fourier transform instruments for that matter, and pass it through an integral transform with a digital processor to deliver a modulation sideband spectrum. We call the result a frequency domain measurement - and don't worry about the fact, other than to make sure we are performing it properly, that the initial data was a time record. References: 18 and 19.

The high isolation power splitter in Figure 13 is worth a moment or two to look at the kind of frequency pulling effect that the two standards at the same frequency, can have on each other. For instance, in Figure 14 consider a quartz crystal resonator inside the standard with a Q of 1 million and 60 dB of net reverse transfer isolation, which would be a typical case for a quartz oscillator.

This would be the case if the signal coming out of the oscillator sees 20 dB of gain then, with 80 dB of gross isolation between the crystal and the output, the net is 60 dB. In this case, the oscillator in question could suffer around 3 parts in 10^{10} of frequency pulling. Clearly this deserves consideration when two standards are close to each other in frequency.

Calibration

When calibrating any system it is desirable to keep the procedure as simple and fool-proof as possible, minimizing the number of dependencies on the calibrations of support instruments. For a phase noise measurement system, the simplest general approach would be to inject a known signal and note the response of the entire system at once. Since many of the devices in a measurement system will be operating at input or output impedance levels substantially departing from 50Ω, with this departure being level dependent, it is highly desirable to require no attenuator setting changes or signal level changes between calibration and measurement. The departures from 50Ω arise unavoidably from several causes: 1. Many high quality signal sources, though designed to drive a 50Ω load, do not present a 50Ω source impedance; 2. Many devices exhibit their best noise performance at impedance levels quite different from maximum power transfer; 3. Inherently non-linear mixers have input impedances which change radically over various portions of the input waveform and are varied further by the level and phase of the other input signal and the terminating impedance.

The simplest high accuracy calibration scheme found to date is based on a single ratio of a pair of RF power measurements at the same port and similar frequencies. This kind of measurement can be performed with 0.1 to 0.5 dB uncertainties depending on frequency, for up to a 90 dB ratio with off-the-shelf standard instruments.

As shown in Figure 15, the calibrating signal is combined with the signal under test from an oscillator or other devices. The calibrating signal may just as well be combined with the signal from the reference source, expecially if this results in a more convenient set-up.

The calibration procedure consists of two parts: First the levels of the calibration signal and the main signal with which it is combined are measured, second, the overall system response to the combined signals is measured. The

computation which combines these three measurements (sometimes with other constants) to yield the scale factor for the frequency domain stability measurement may be considered a third step.

Since the details of this procedure can affect the resulting accuracy, the following sequence is suggested:

1. Connect calibration signal source to combiner. Terminate other input of combiner. Connect output of combiner to power meter. Set level of calibration signal source, as shown on power meter, to desired level, at least 40 dB below the output level of the source under test. Measure C volts rms. These levels should have been pre-determined by preliminary measurements. This should likely consist of several passes through the entire calibration and measurement sequence to establish workable and convenient signal levels and control settings.

2. Connect oscillator under test in place of termination of input of power combiner. On power meter, measure M volts rms.

3. Disconnect power meter from output of combiner and connect combiner to input of measurement system.

4. Read measurement system response at pseudo-sideband modulation frequency corresponding to the difference between frequency of C signal and M signal.

5. Compute measurement system scale factor based on reading in 4. and the fact that the input to the measurement system has a peak phase deviation of C/M radians.

6. Disconnect calibration signal source from input to combiner and terminate combiner input. Calibration is now complete. Proceed with measurements;

Other than errors in the power meter's calibration, (on a ratio basis only, absolute calibration being of no consequence) the only other error source in this calibration signal set-up can arise from non-50 Ω (or in general, unmatched) impedances. This concern arises only if the input

port of the measurement system differs from the power meter in impedance. This would almost always be true if a mixer is the input element of the system.

There is still no problem unless there is also a difference in the source impedance appearing at the output of the combiner for the calibration signal versus the main signal. This would again be true very typically for a frequency standard output which is designed to be loaded with 50 Ω but does not provde a 50 Ω source impedance. The calibration signal source is likely to be a signal generator whose output source impedance (probably through an attenuator) is a fairly accurate 50 Ω . These signals may pass through a directional coupler which would tend to normalize any mismatch of the calibration source, while presenting the main signal source impedance essentially unchanged. Should this error source be present, its effect can be calculated, and/or impedance matching can alleviate it. Figure 16 shows alternatives for the coupler. Figure 17 is a vector picture of what is going on; in sine wave terms, the calibrate signal is a vector of a slightly different frequency from the main signal, and the vector resultant is both AM and PM. The system supposedly responds only to the PM, in the cases being studied here. For a very small calibrate signal and a large main signal, the phase excursion is very well defined. There is a form following Figure 17 to help keep track of the arithmetic in this calibration and measurement scheme.

Analog Analyzers

When using an analog wave or spectrum analyzer to measure the spectrum of the output of a system, it is necessary to be very cautious of the actual noise bandwidth of the instrument. The resolution bandwidth at the switch on the front panel, when set to 10 Hz, or 3 Hz, or 1 Hz will not be equal to the noise bandwidth. This can, in most cases, account for on the order of 1 dB of measurement error. For a crude measurement, this may be of no concern. Usually, the noise bandwidth is wider than the resolution bandwidth on the front panel. The number of poles in the IF that determines that resolution bandwidth versus the slope of the noise can cause problems too. This is well covered in reference: 11.

When averaging follows a log amplifier function, about 2.5 dB of error occurs due to skewing of the mean, see reference: 10. In some of the analog instruments which utilize digital storage, there is a circuit to catch the

peak of the bright lines as they sweep past them. When noise is measured with an analog to digital converter circuit which operates in that way, it catches the peaks of the noise and gives an answer that is higher by as much as 6 dB, and in most cases, 2-4 dB, than the true noise level.

Digital Analyzers

In the digital Fourier transform or fast Fourier class of analyzers, we have had a good deal less experience. When using these analyzers, a point to watch for is the setting of the knob which can be switched to sine or random. The software takes care of the problems of logging and noise bandwidth, if that switch is in the right position. When a sinewave calibration is being made, that switch should be in the sine position. Then, when noise is being measured, the switch must be in the random position. However, for discrete spur measurement, the sine position must be used. Some problems have arisen in trying to utilize these analyzers because their gain is insufficient to measure the small signals coming out of the mixers. Measuring down below 1 Hz, for the modulation frequency, is the vast step forward that these analyzers are offering. Yet, in some cases, a very dc stable or low frequency stable amplifier is needed between the mixer and the analyzer itself. This amplifier must also have good noise performance, usually a trade-off with long-term stability. It seems well to be cautious about the performance of such an amplifier.

REFERENCES

1. Winkler, G.M.R., "A Brief Review of Frequency Stability Measures", 8th PTTI, 1976, pp.489-527.

2. Rutman, Jacques, "Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress", Proc. IEEE, Vol. 66, No. 9, September 1978, pp. 1048-1075.

3. Barnes, James A., et al, "Characterization of Frequency Stability", IEEE Trans. on Instrumentation and Measurement, Vol. IM-20, No. 2, May 1971, pp. 105-120.

4. Ashley, J. Robert, et al, "The Measurement of Noise in Microwave Transmitters", IEEE Trans. on Microwave Theory and Techniques, Vol. MTT-25, No. 4, April 1977, pp. 294-318.

5. Lance, A. L., et al, "Phase Noise Characteristics of Frequency Sources", 9th PTTI, 1977, pp. 463-483.

6. Parzen, Benjamin and Hou, James, "A Technique for a Self-Phase Noise Measuring System for Signal Sources", 32nd FCS, 1978, pp. 432-438.

7. Motchenbacher, C. D. and Fitchen, F. C., Low-Noise Electronic Design, Wiley, New York, 1973.

8. Caruthers, R. S., "Copper Oxide Modulators in Carrier Telephone Systems", Bell System Technical Journal, Vol. 18, No. 2, April, 1939, pp. 315-337.

9. Kurtz, Stephan R., "Specifying Mixers as Phase Detectors", Microwaves, Vol. 17, No. 1, Jan. 1978, pp. 80-89.

10. Fischer, Michael C., "Frequency Stability Measurement Procedures", 8th PTTI, 1976, pp. 575-617.

11. Walls, F. L. and Stein, S. R., "Accurate Measurements of Spectral Density of Phase Noise in Devices", 31st FCS, 1977, pp. 335-343.

12. Lance, A. L., et al, "Phase Noise Measurements Using Cross Spectrum Analysis", Conference on Precision Electromagnetic Measurements Digest, 1978, pp. 94-96.

13. Hewlett-Packard Staff, "Understanding and Measuring Phase Noise in the Frequency Domain", Application Note 207, 1976.

14. Lance, A. L. et al, "Automatic Phase Noise Measurements in the Frequency Domain", 31st FCS, 1977, pp. 347-358.

15. Walls, F. L., et al, "Design Considerations in State-of-the-Art Signal Processing and Phase Noise Measurement Systems", 30th FCS, 1976, pp. 269-274.

16. Peregrino, Luiz, and Ricci, David W., "Phase Noise Measurement Using A High Resolution Counter with On-Line Data Processing", 30th FCS, 1976, pp. 309-317.

17. Hewlett-Packard Data Sheet "5390A Frequency Stability Analyzer" 1978.

18. Allan, David W. "Report on NBS Dual Mixer Time Difference System (DMTD) Built for Time Domain Measurements Associated with Phase 1 of GPS", NBSIR 75-827, National Bureau of Standards, Boulder, Co. 80302, 1976.

19. Hewlett-Packard Data Sheet "5390A Frequency Stability Analyzer", Option 10, 1978.

Abbreviations Used In References:

PTTI: Precision Time and Time Interval Applications and Planning Meeting Proceedings; Technical Information & Administrative Support Division, Code 250, Goddard Space Flight Center, Greenbelt, Md. 20771, Telephone (301) 982-4488.

FCS: Frequency Control Symposium Proceedings; 1976: National Technical Information Service ADA046089

1977 and 78: Electronic Industries Association 2001 Eye Street NW, Washington, D.C. 20006.

Table 1
FREQUENCY DOMAIN MEASUREMENT SYSTEMS

1. Various Measures

2. Various Devices to be Tested

3. Measurement Methods and Hook-Ups

   - Applications

   - Calibration

   - Traps

4. Frequency Domain Analyzer Instruments

   - Ranges, Accuracy

   - Traps

Table 2
FREQUENCY DOMAIN STABILITY MEASURES

MEASURE	SYMBOL	UNITS
† Spectral density of of fractional frequency fluctuation	S_y	1/Hz
Spectral density of frequency fluctuation	$S_{\Delta f}$	Hz^2/Hz
† Spectral density of phase fluctuation	S_ϕ (also $S_{\delta\phi}$, $S_{\Delta\phi}$)	rad^2/Hz
* Single sideband phase noise to carrier ratio	$\mathcal{L}$	1/Hz
Residual frequency modulation, rms, in bandwidth f_B, located f_m from carrier	Δf_{res}	Hz
Residual phase modulation, rms, in modulation spectrum between f_1 and f_2	$\Delta\phi$	radians

† Recommended by IEEE, CCIR
* Widely used on procurement specifications and data sheets.

Independent Variable: Frequency

Modulation Frequency	f_m
Fourier Frequency	f
Sideband Frequency	f
Offset Frequency	f
Baseband Frequency	f

Figure 1
SIMPLEST FREQUENCY DOMAIN MEASUREMENT SYSTEM

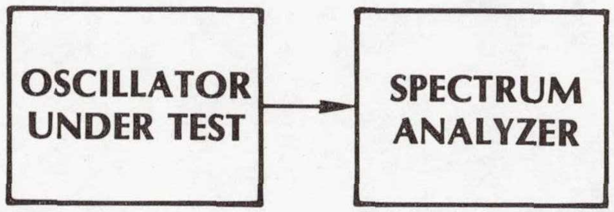


Figure 2
FREQUENCY DOMAIN MEASUREMENT SYSTEM

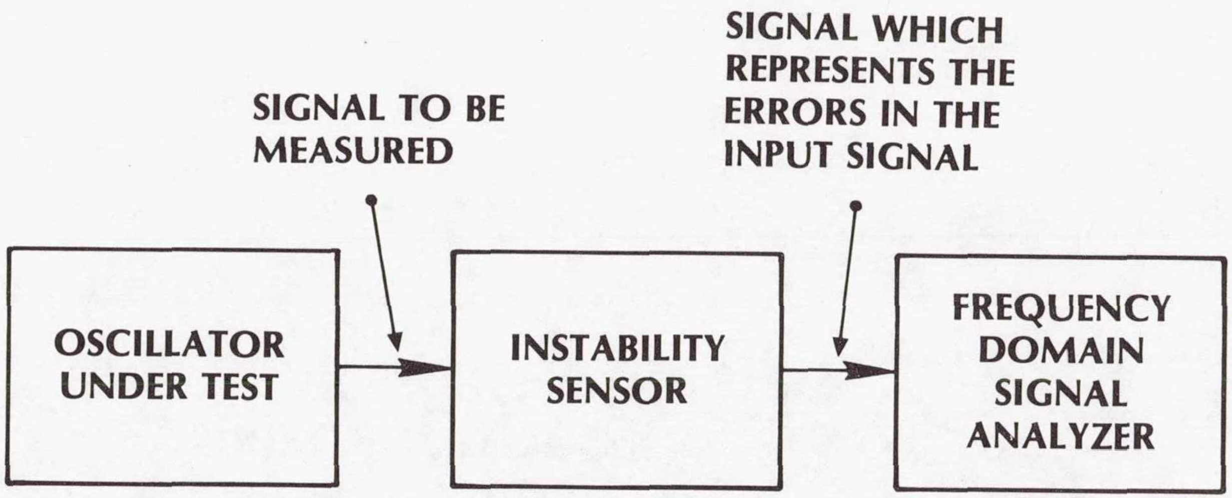


Figure 3
AUTOCORRELATION DEMODULATOR

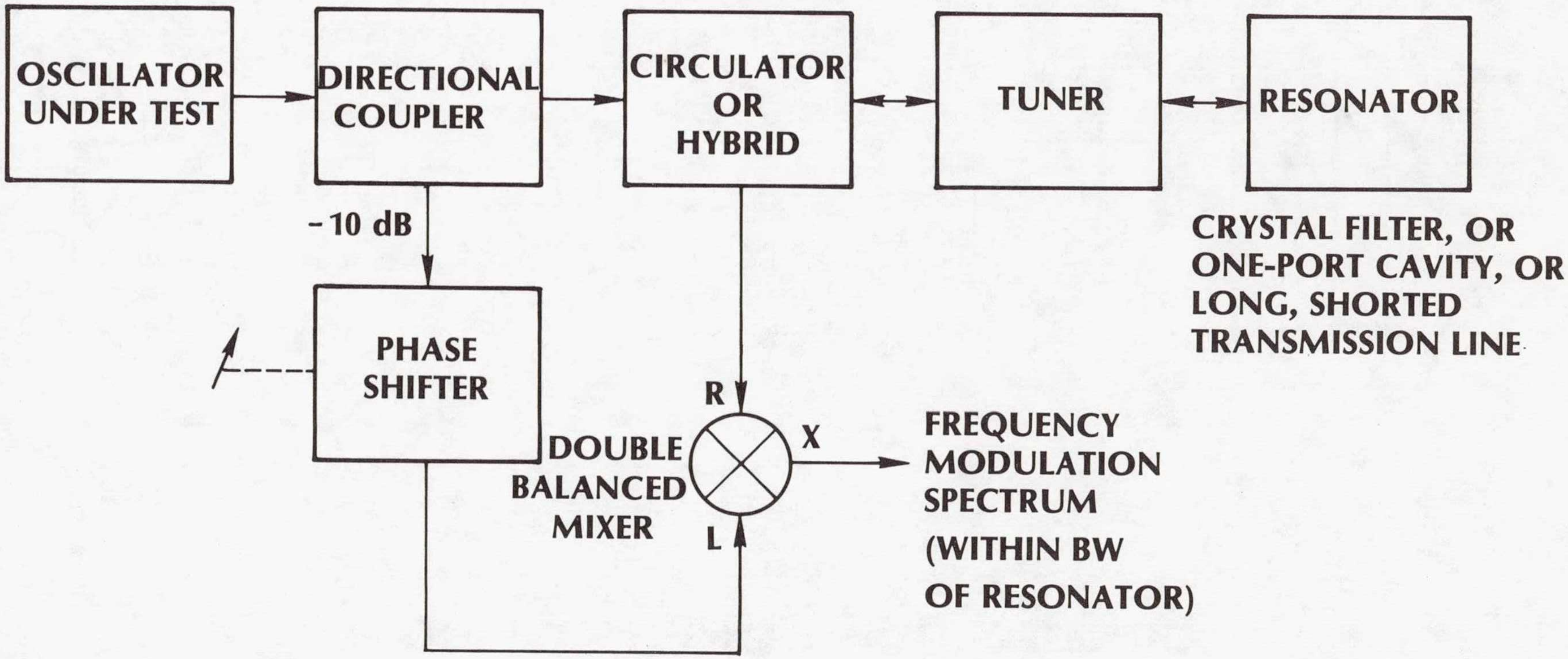


Figure 4

DOUBLE BALANCED MIXER OPERATION

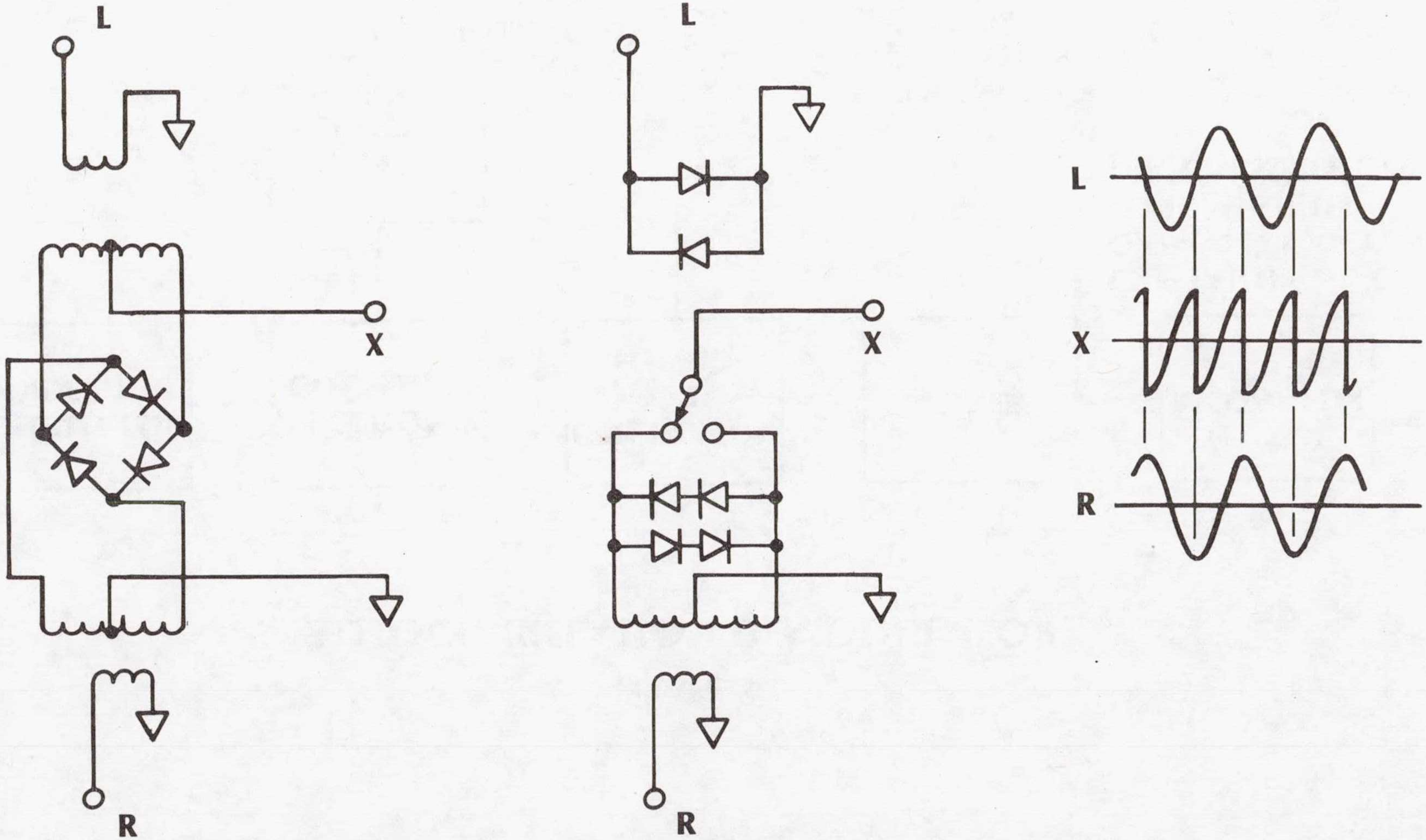
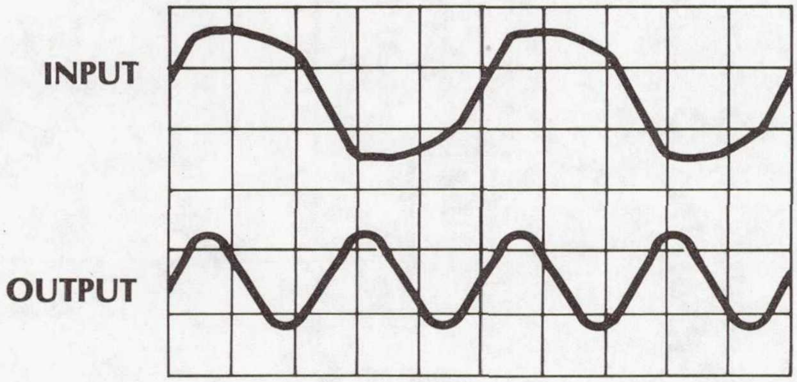


Figure 5
DOUBLE BALANCED MIXER PHASE DETECTOR WAVEFORMS AND SPECTRUM

0.5 V/div
10 MHz
TO L (R same)
(+7 dBm incident)

0.5 V/div
20 MHz
From X



10 dB/div
10 MHz/div
Spectrum of above output from X

Top of grat. = +7 dBm

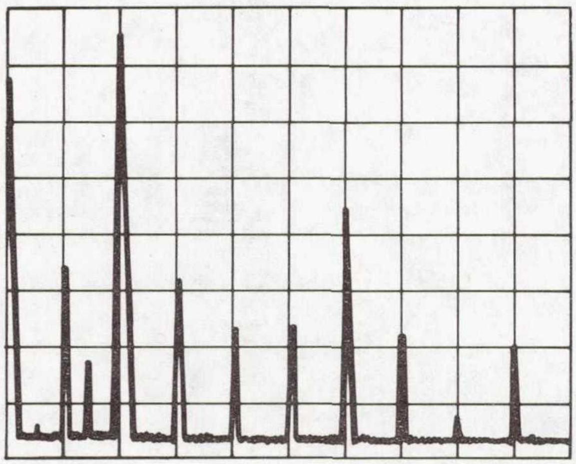


Figure 6
SIMPLE LOW PASS MIXER TERMINATION

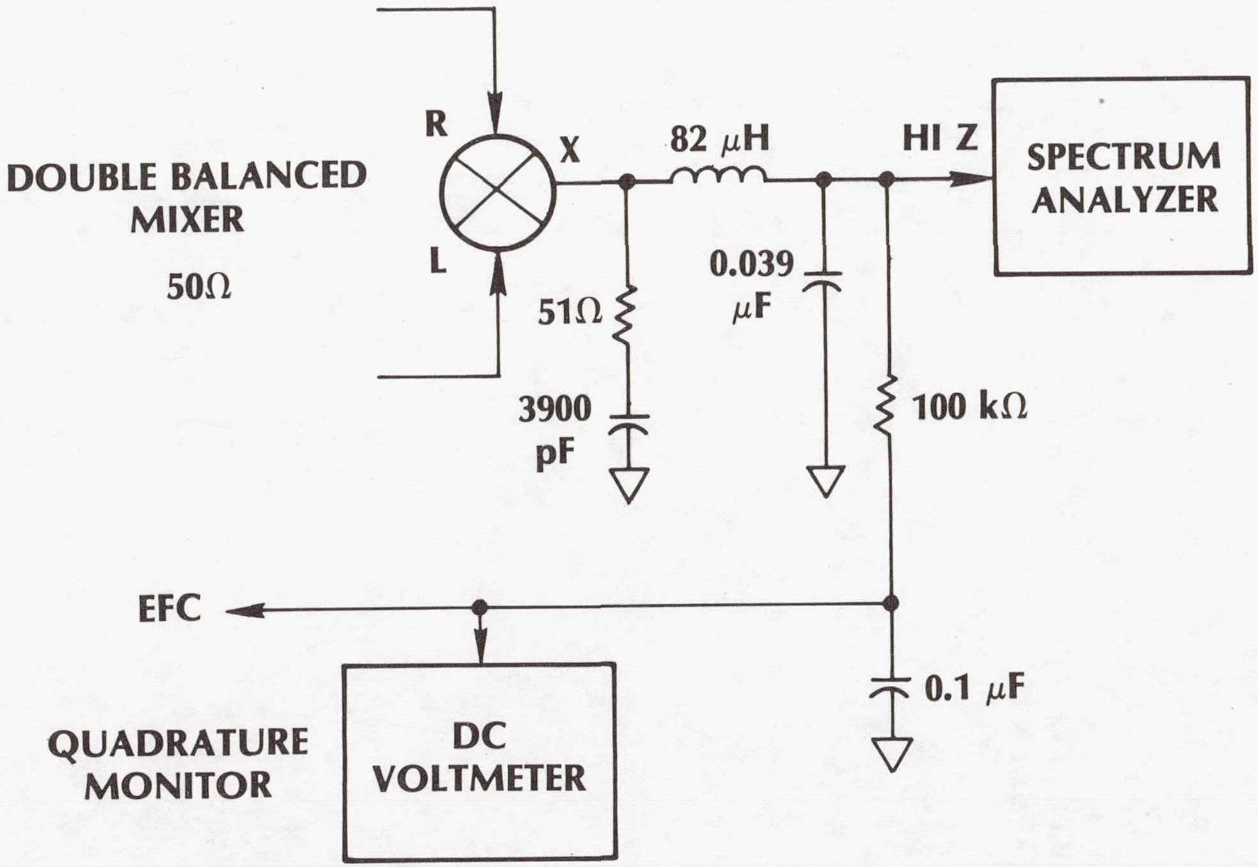
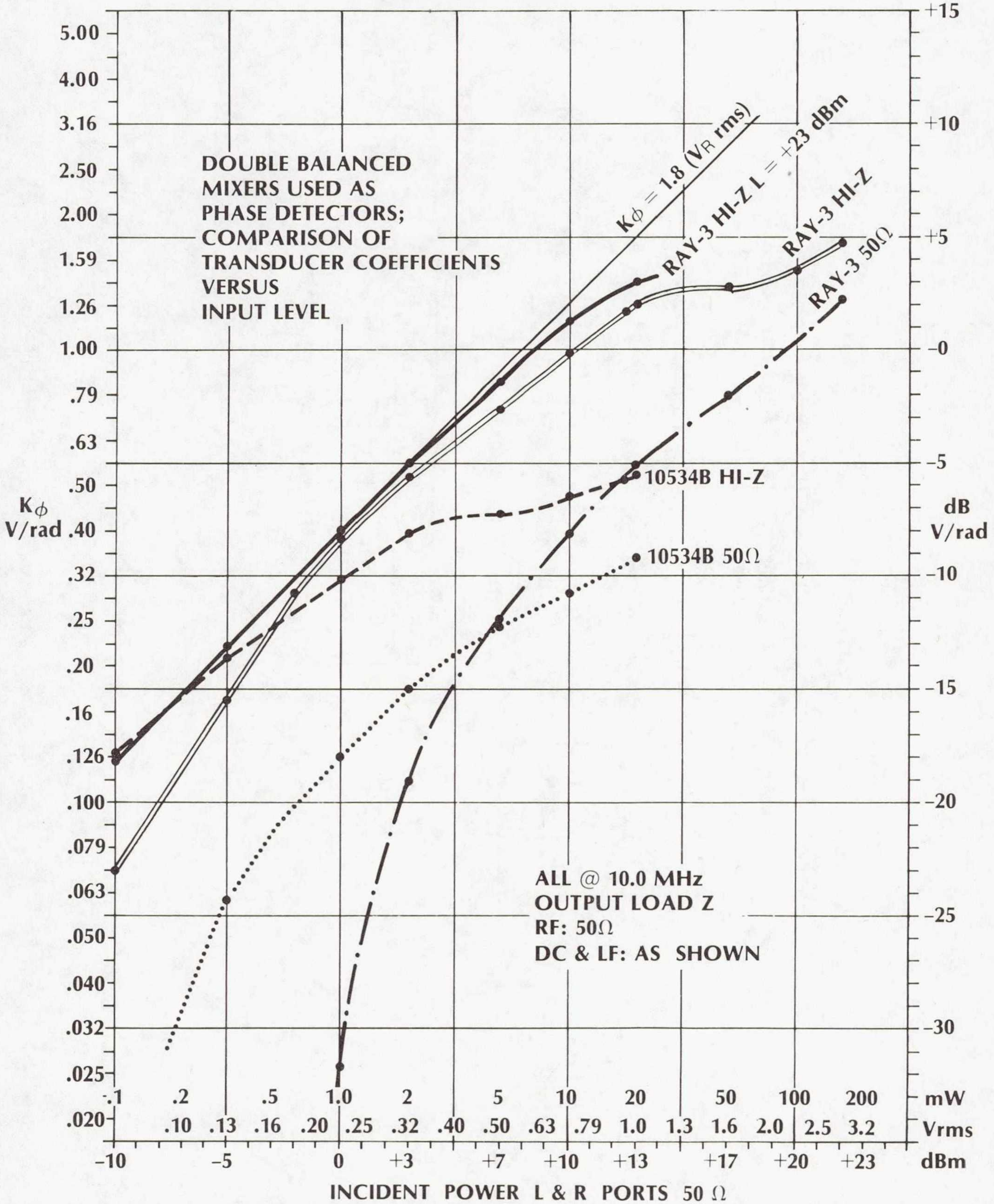


Figure 7



DOUBLE BALANCED
MIXERS USED AS
PHASE DETECTORS;
COMPARISON OF
TRANSDUCER COEFFICIENTS
VERSUS
INPUT LEVEL
Kϕ = 1.8 (VR rms)
RAY-3 HI-Z L = +23 dBm
RAY-3 HI-Z
RAY-3 50Ω
10534B HI-Z
10534B 50Ω
ALL @ 10.0 MHz
OUTPUT LOAD Z
RF: 50Ω
DC & LF: AS SHOWN
Kϕ V/rad
5.00
4.00
3.16
2.50
2.00
1.59
1.26
1.00
.79
.63
.50
.40
.32
.25
.20
.16
.126
.100
.079
.063
.050
.040
.032
.025
.020
+15
+10
+5
−0
−5
−10
−15
−20
−25
−30
dB V/rad
.1 .2 .5 1.0 2 5 10 20 50 100 200 mW
.10 .13 .16 .20 .25 .32 .40 .50 .63 .79 1.0 1.3 1.6 2.0 2.5 3.2 Vrms
−10 −5 0 +3 +7 +10 +13 +17 +20 +23 dBm
INCIDENT POWER L & R PORTS 50 Ω

Figure 8
DUAL AUTOCORRELATOR, CROSS CORRELATION ENHANCED

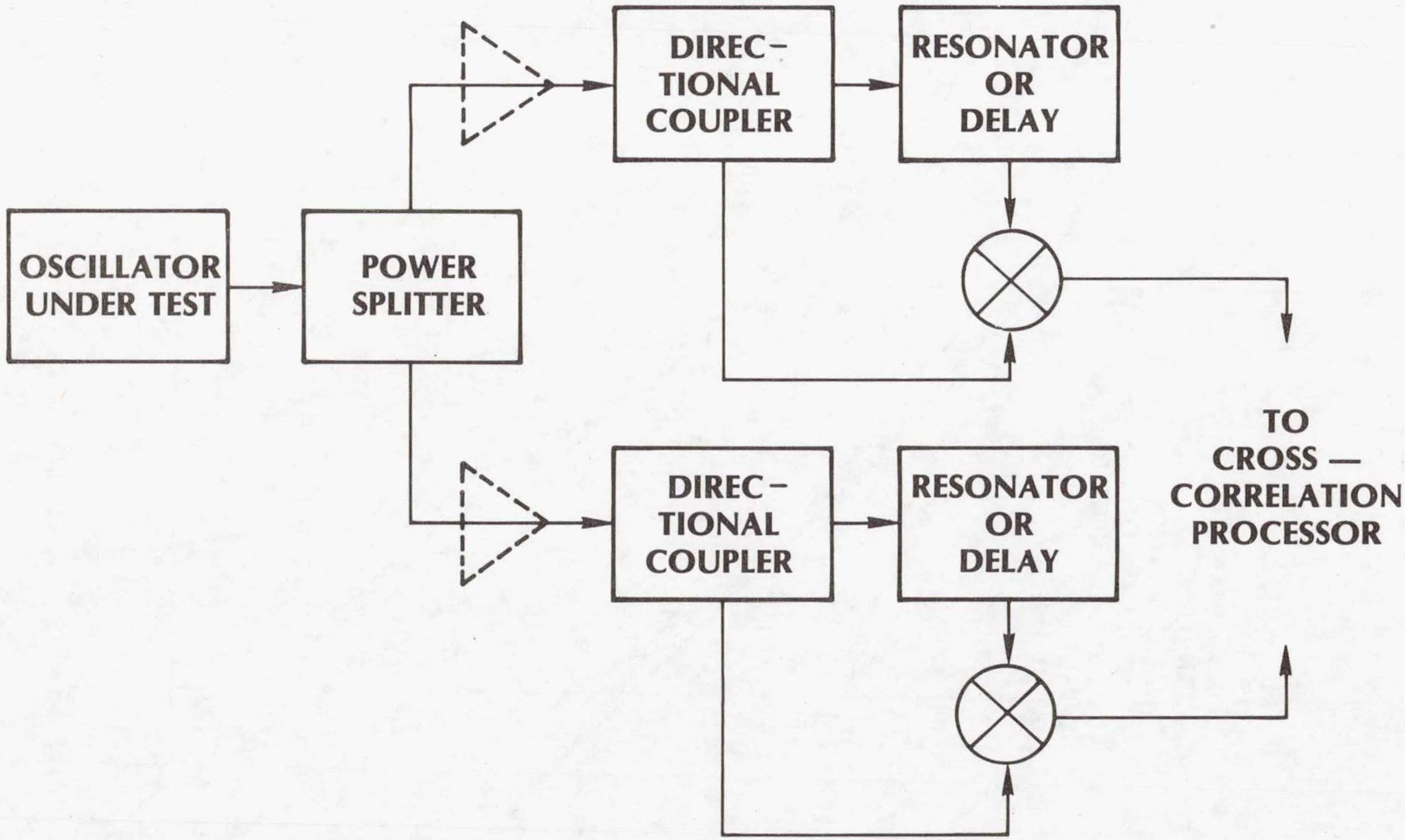


Figure 9

TWO OSCILLATOR, LOCKED

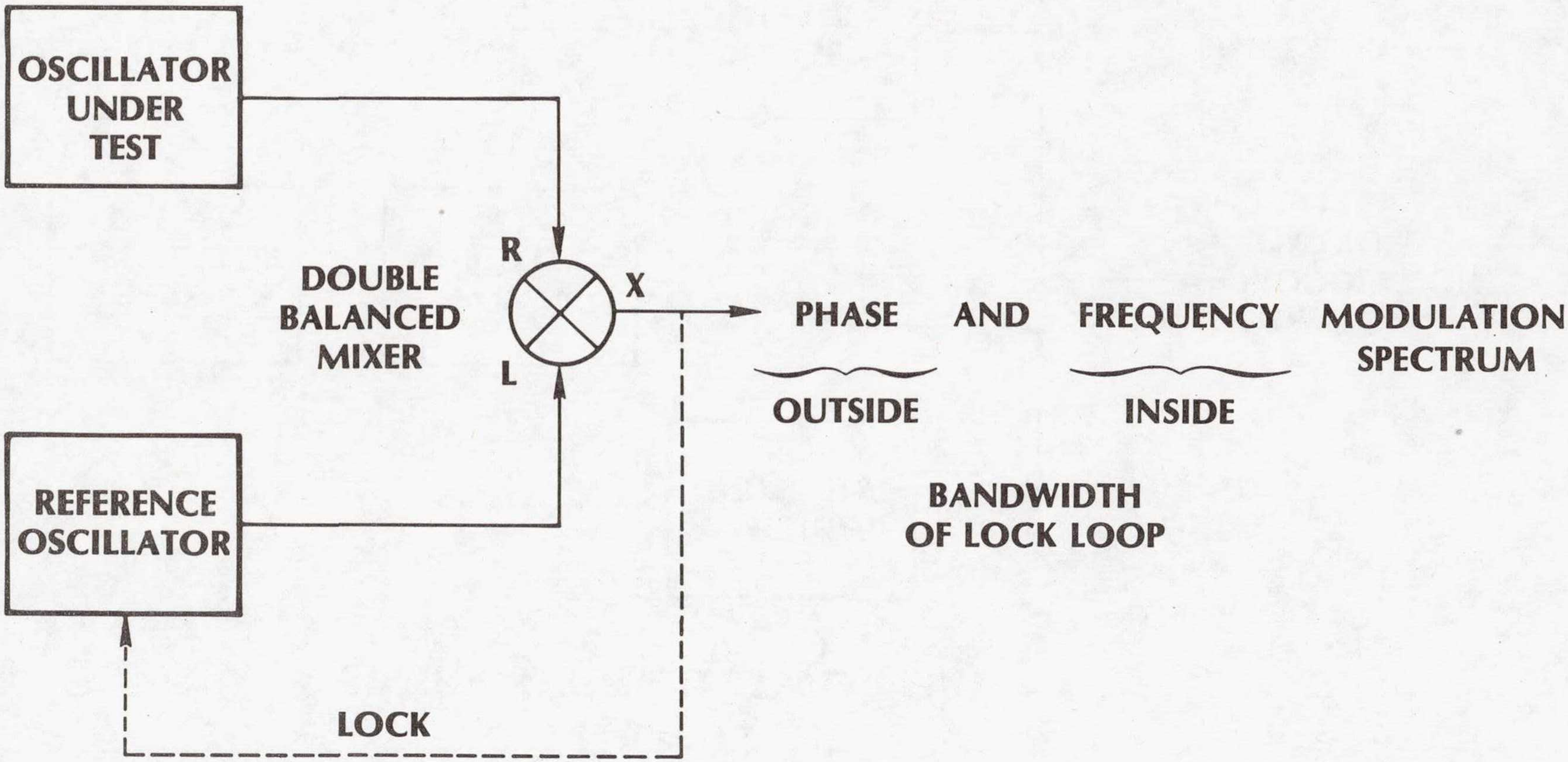


Figure 10

SINGLE OR TWO, TWO-PORT

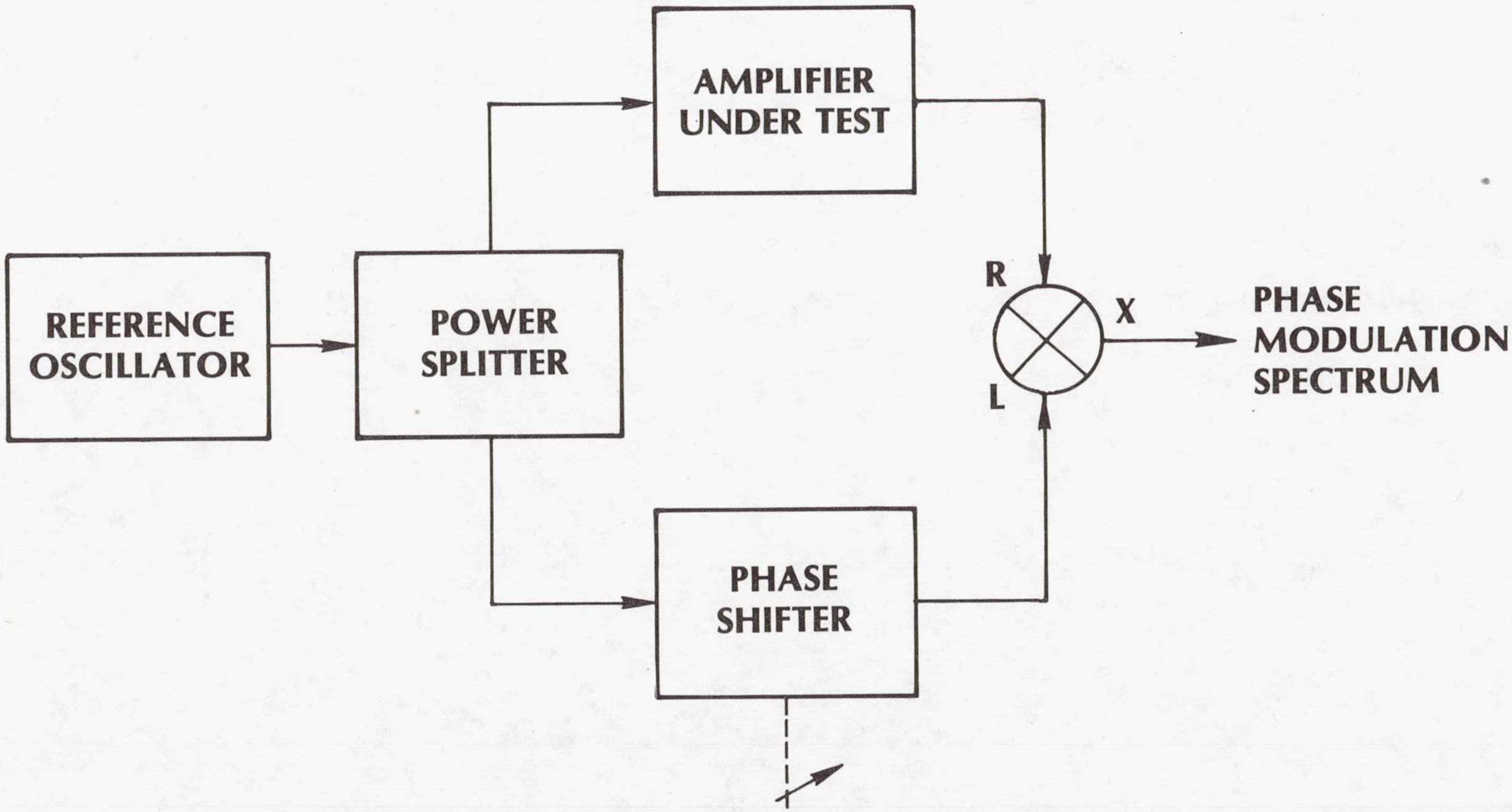


Figure 11
TWO OSCILLATOR, LOCKED, DUAL MIXER, CROSS CORRELATION ENHANCED

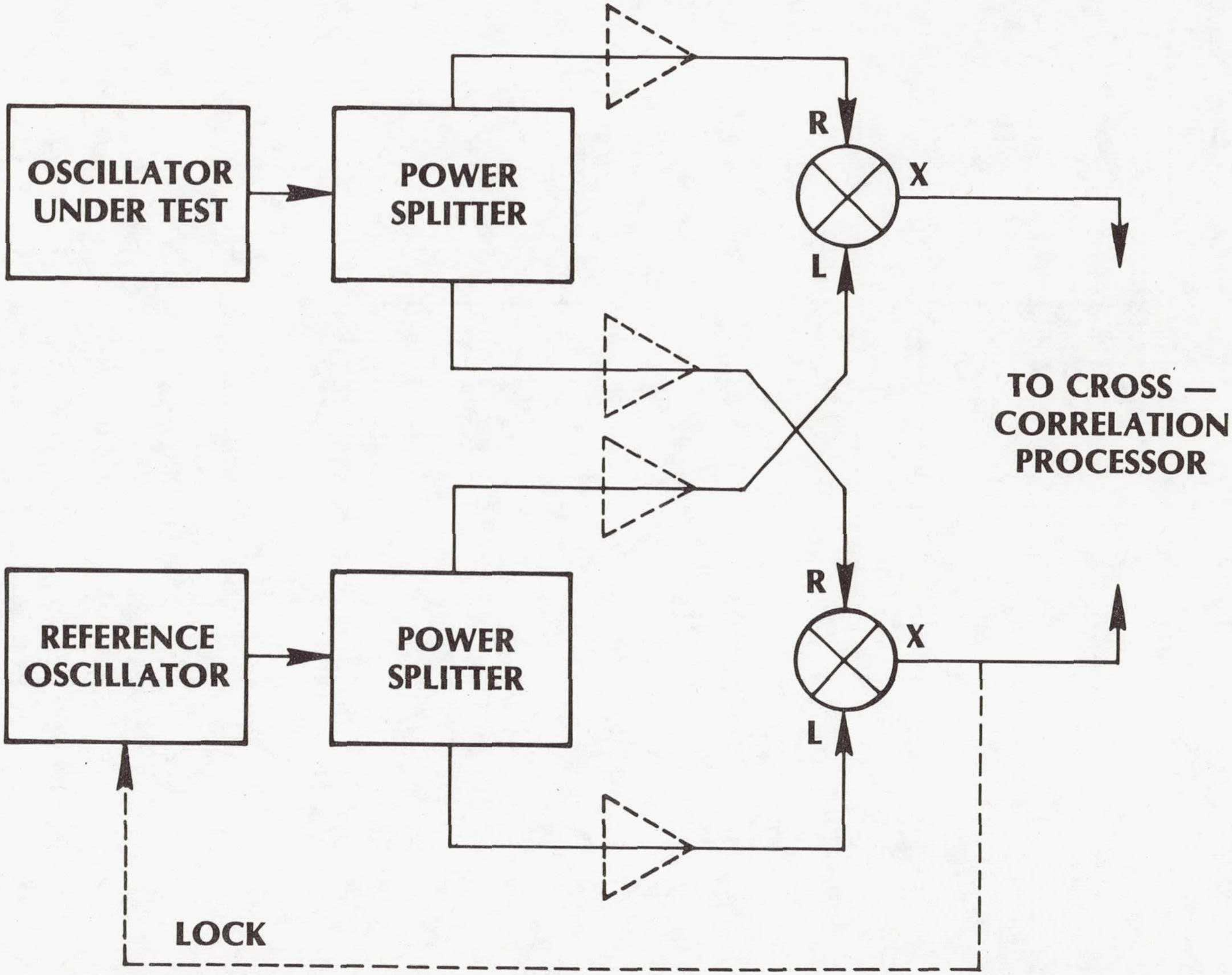
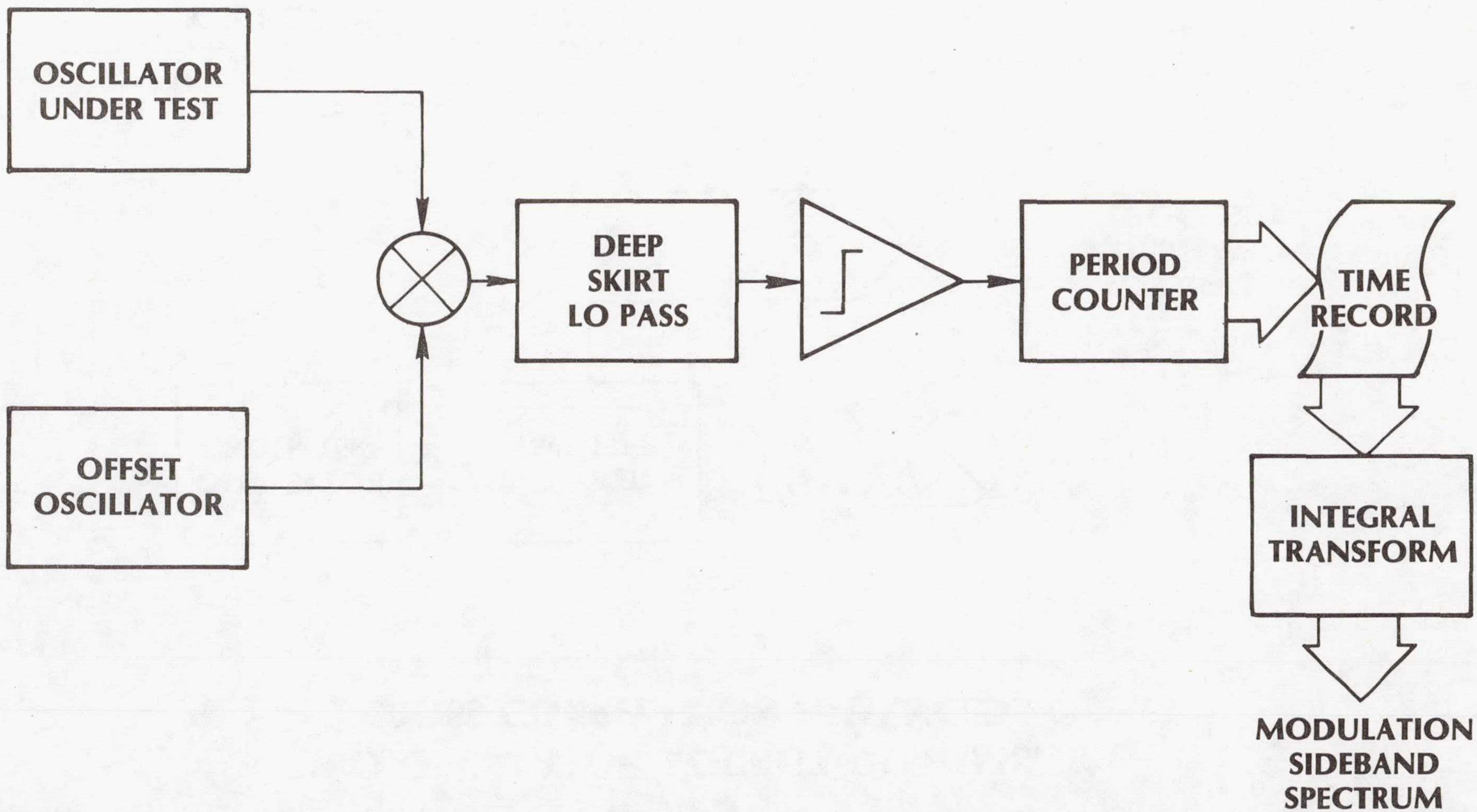


Figure 12
TWO OSCILLATOR, OFFSET



370

Figure 13
THREE OSCILLATOR, TWO NOT OFFSET DUAL MIXER TIME DIFFERENCE

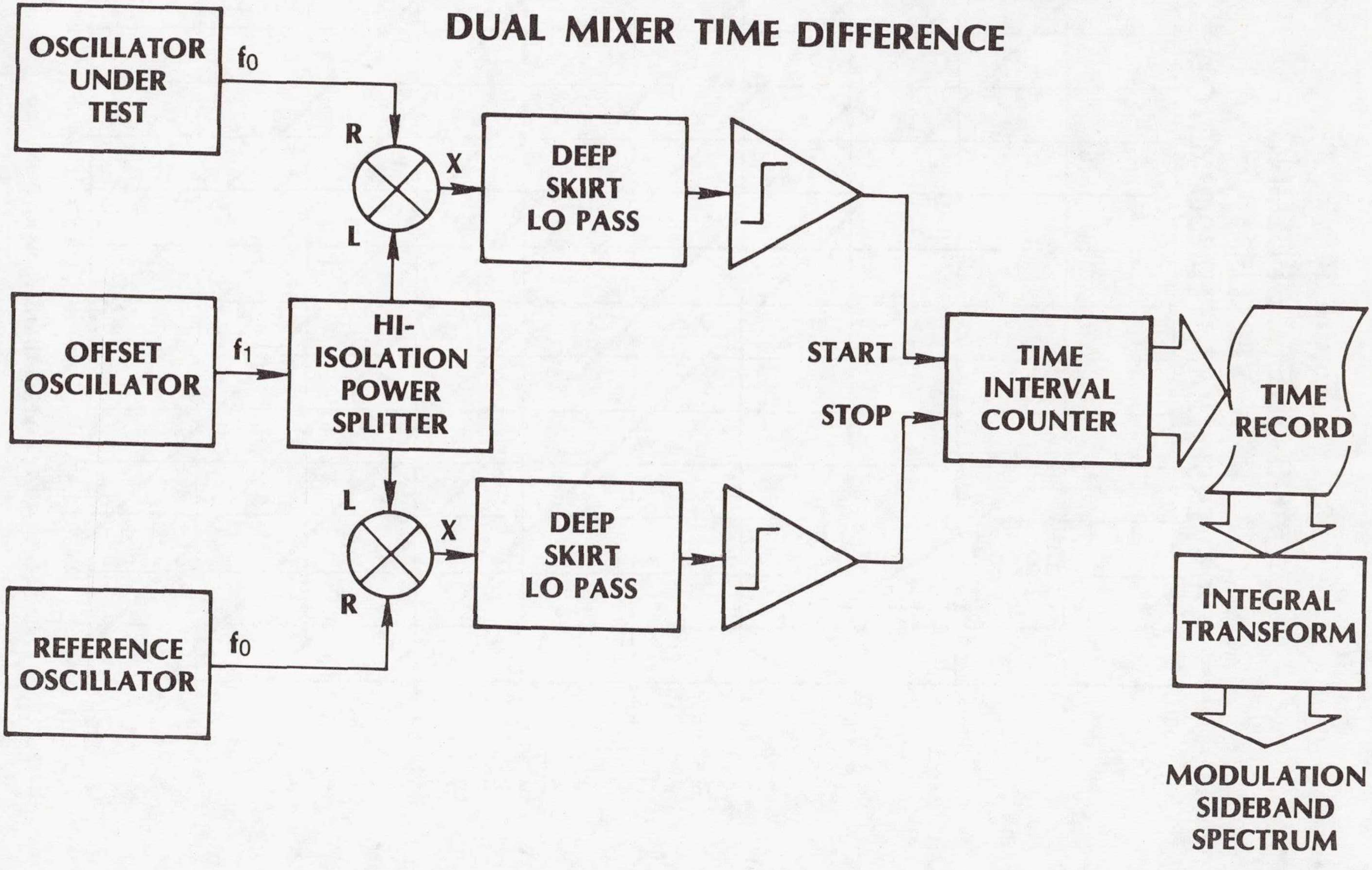


Figure 14
FREQUENCY PULLING VERSUS REVERSE TRANSFER ISOLATION

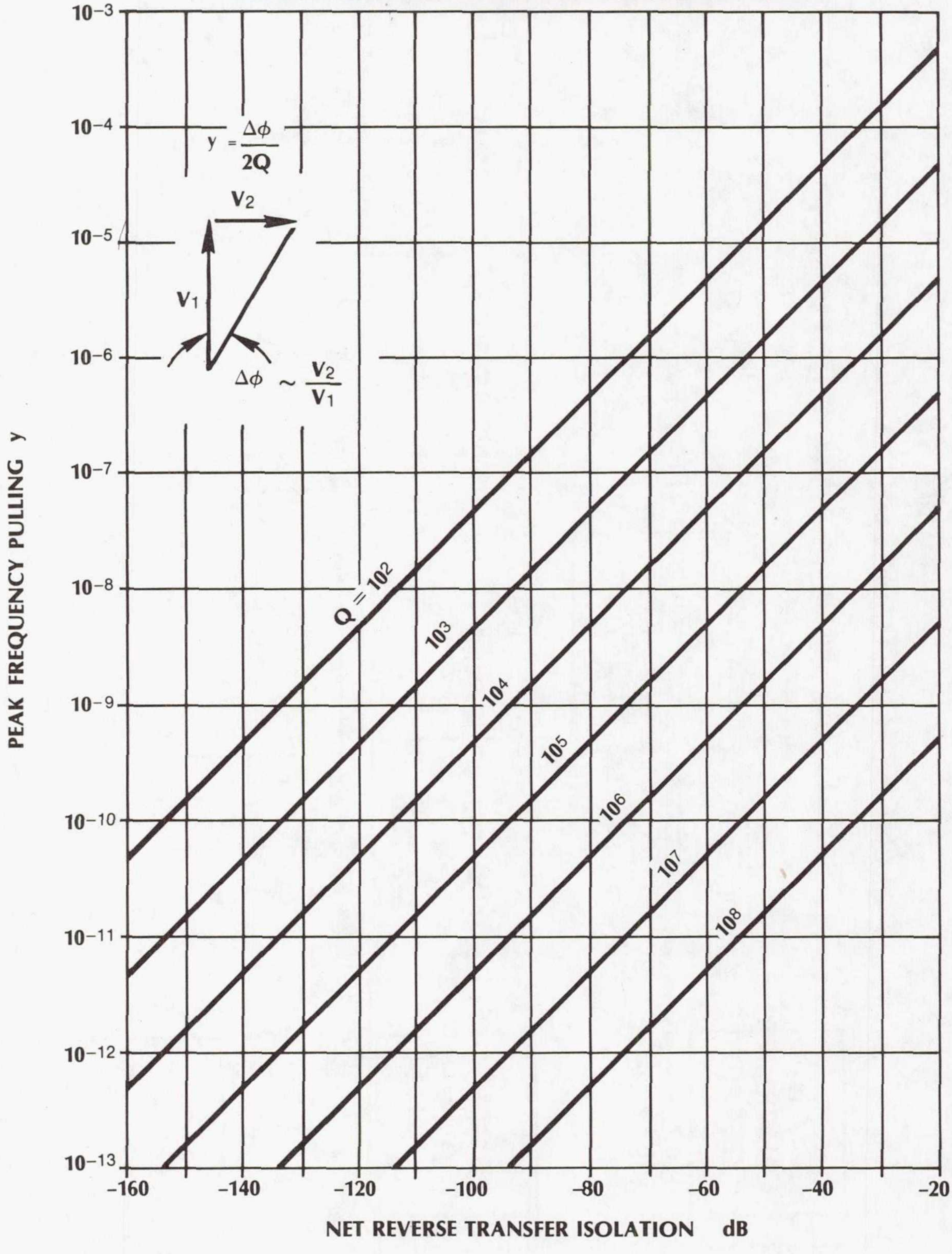


Figure 15
DIRECT RATIO CALIBRATION METHOD

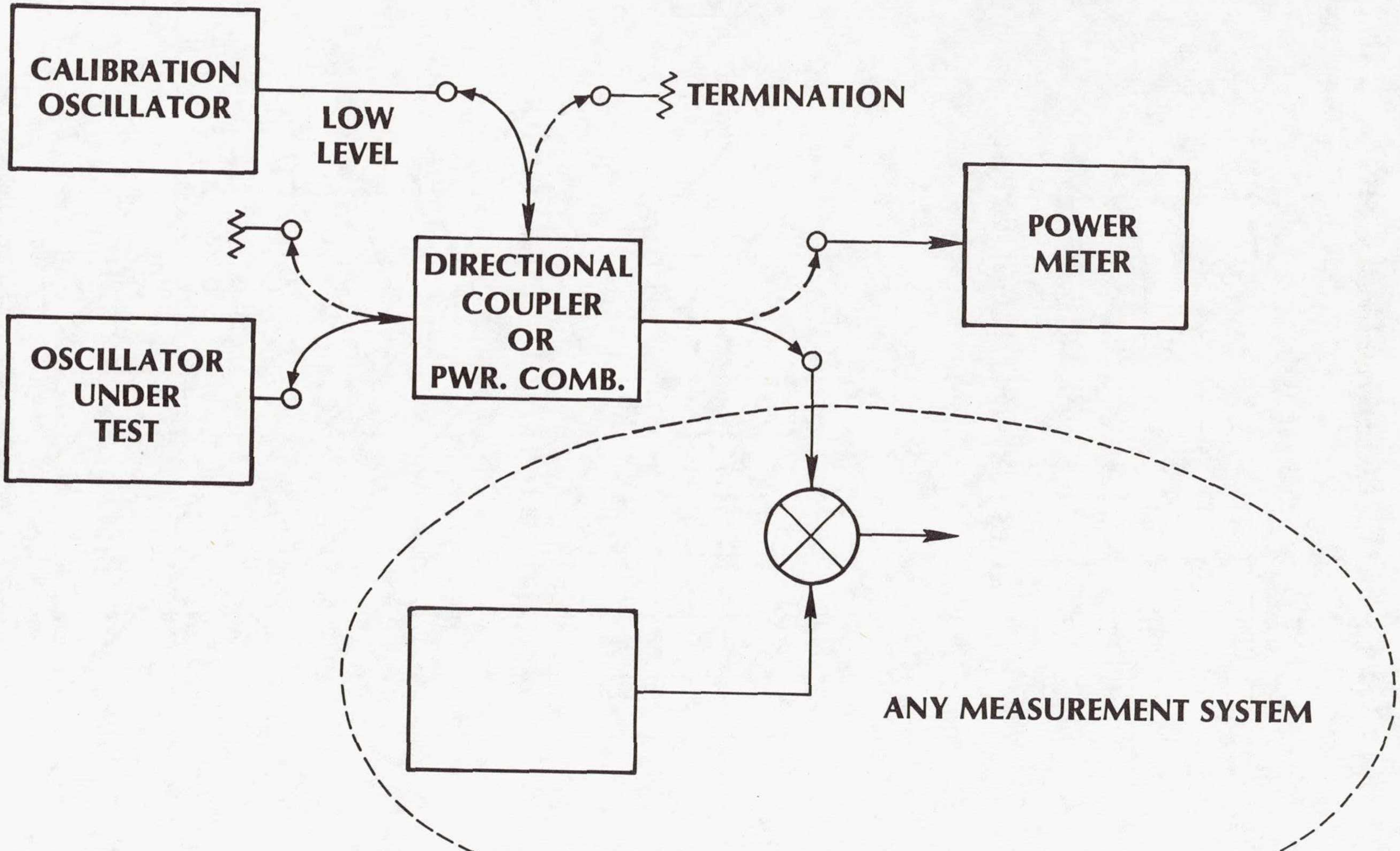
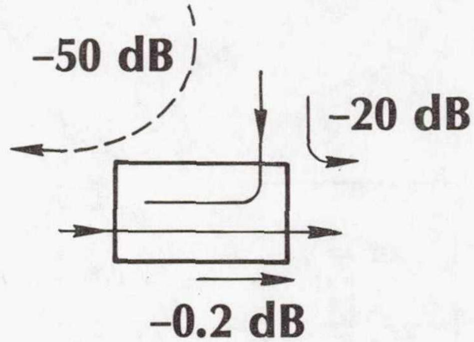
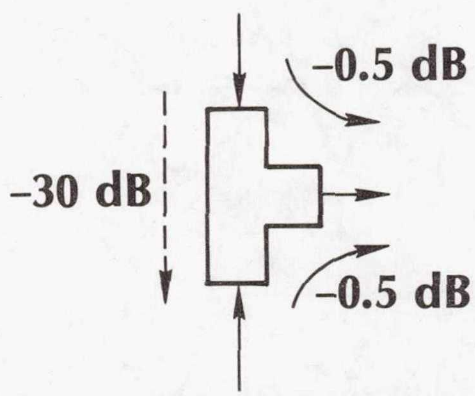


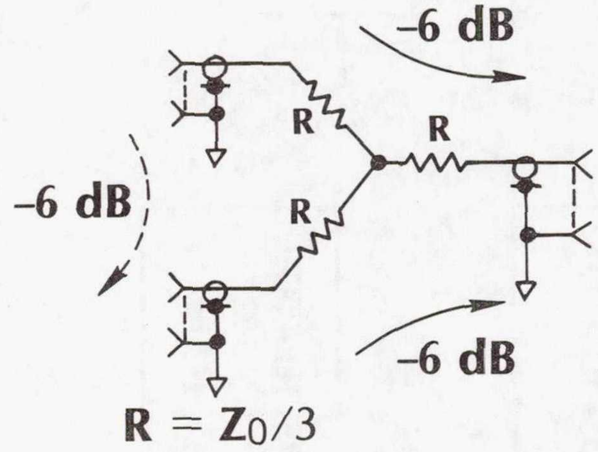
Figure 16
ALTERNATIVES FOR CALIBRATION SIGNAL INJECTION



20 dB DIRECTIONAL COUPLER



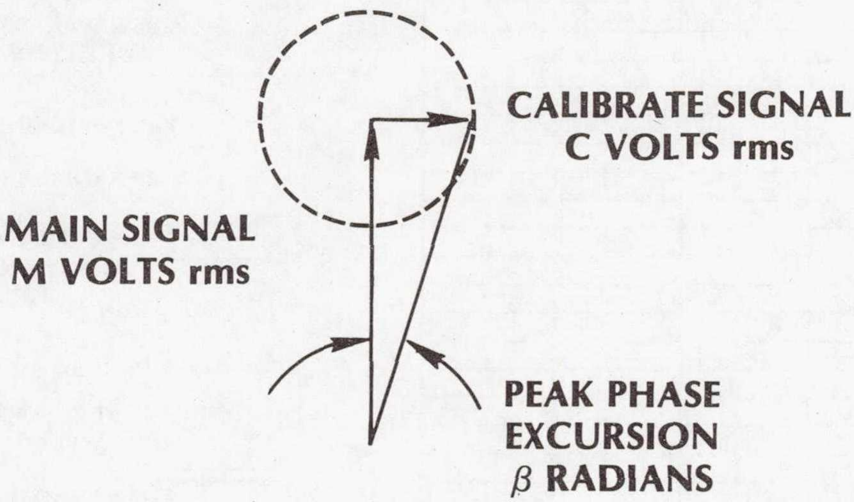
3 dB POWER SPLITTER-COMBINER



$R = Z_0/3$

6 dB POWER SPLITTER-COMBINER

Figure 17
CALIBRATION VECTOR DIAGRAM



$$\beta = \arcsin \frac{C}{M}$$

For small angles, $\beta << 1$

$$\frac{C}{M} \sim \beta \equiv \text{Modulation index, radians, peak}$$

SINGLE SIDEBAND PHASE NOISE ($\mathcal{L}(f)$) MEASUREMENT WORKSHEET

Using a Double-Balanced Mixer, Directional Coupler and Spectrum Analyzer

L port signal level should be the maximum available, up to the mixer specs, and must remain constant for both calibrate and measure.

FREQ. Hz	RAW READOUT dB	− SYST. RESP. dB	+ SCALE FACTOR dB	= MEAS. DATA dB
1.0				
1.3				
1.6				
2.0				
2.5				
3.2				
4.0				
5.0				
6.3				
8.0				
10				
13				
16				
20				
25				
32				
40				
50				
63				
80				
100				
130				
160				
200				
250				
320				
400				
500				
630				
800				
1.0k				
1.3k				
1.6k				
2.0k				
2.5k				
3.2k				
4.0k				
5.0k				
6.3k				
8.0k				
10k				
13k				
16k				
20k				
25k				
32k				
40k				
50k				
63k				
80k				
100k				

TEST CONDITIONS AND SCALE FACTOR COMPUTATION

LO SIG INTO L PORT (   ) dBm

CALIBRATE FREQ. (   ) Hz

CALIBRATE

− (   ) dBm MAIN SIG. } **INTO R PORT**
+ (   ) dBm CAL SIG.

− (−   ) dB INPUT SENS.

+ (−   ) dB SYSTEM RESP.
@ CAL FREQ.

− (−   ) dB RAW CAL READOUT
"ADJUST" LAMP OUT?

MEASURE

+ (−   ) dB INPUT SENS.

−6 dB RADIANS TO SSB

$\boxed{}$ **dB BRIGHT LINE SCALE FACTOR**

"ADJUST" LAMP? (   )

− (   ) dB=10 log (Hz BW)

+1.7 dB $\left(\begin{array}{l}+\ 2.5 \text{ logging} \\ -\ 0.8 \text{ Gauss BPF}\end{array}\right)$

$\boxed{}$ **dB RANDOM NOISE SCALE FACTOR**

SMOOTHING? OR VIDEO FILTER? (   )

UNIT UNDER TEST: ______________________

REFERENCE SIGNAL SOURCE: ______________________

MIXER: ______________________

AC AMP: ______________________

DC AMP: ______________________

NAME: ______________________

DATE: ______________________

QUESTIONS AND ANSWERS

DR. MICHEL TETU, Laval University:

I would like to point out first that when you use digital spectrum analyzers, we again have some difficulties in defining the equivalent bandwidth over which the measurement is done. The source seems to be the digital filtering used to compute the spectrum. Second, when you do not have a wide phase noise or a wide noise observed, it is a little dangerous to use the equivalent bandwidth principle.

MR. FISCHER:

I think we are somewhat at the mercy of the vendors of those devices, until we get more familiar with them, at least, to properly characterize their noise bandwidth in a believable way.

Page intentionally left blank

RUBIDIUM FREQUENCY STANDARD TEST PROGRAM FOR NAVSTAR GPS

Frank K. Koide and Darrell J. Dederich
Electronic Systems Group
Rockwell International, Anaheim, California

ABSTRACT

Space-qualified Rubidium Frequency Standards (RFS) are being developed for the NAVSTAR Global Positioning System (GPS) program by Rockwell International Electronic Systems Group.

This paper will present the test data of the RFS Program in the Production Phase and computer automation as an essential element in the evaluation of the RFS performance in a simulated spacecraft environment.

Typical production test data will be discussed for stabilities from 1 to 100,000 seconds averaging time and simulated time error accumulation test. Also, design considerations in developing the RFS test systems for the acceptance test in production will be discussed in the paper. Finally, as part of a life-cycle test in vacuum, stability data to 1,000,000 seconds averaging time will be presented.

INTRODUCTION

The Electronic Systems Group (ESG) of Rockwell International has been involved in the engineering development and production of the Rubidium Frequency Standard for the GPS Phase I effort since 1974. In Phase I of the program, ESG will produce a total of 29 RFS; one prototype, three engineering models, and 25 production units. Currently, all but four units have been delivered to our sister Group, the Space Systems Group who is the Prime Contractor for the satellite NAVSTAR-GPS Program.

The first series of four GPS satellites are in orbit, with each carrying three redundant RFS. Three of the satellites have been declared operational, and initial test results proved that the user equipment is capable of navigation accuracies better than three meters in three dimensions. When the fourth NAVSTAR satellite is declared operational, the prime objective of validating the Global Positioning System, will be accomplished.

The RFS test cycle (Figure 4) covers two major phases, the pre-production and the acceptance level testing. The pre-production testing covers the board and system assembly, where tests are made of the components, absolute frequency, temperature compensation, and repairs, if required. The acceptance-level testing includes the environmental and certification tests. All of the certification tests and part of the pre-production and environmental tests are performed on three test systems located within the Metrology Lab-Laboratory of ESG.

Testing completed to date indicates that the RFS meets or exceeds the GPS specifications for Phase I, and is a potential candidate for the forthcoming Phase II/III Programs.

TEST SYSTEM DESCRIPTION

The three automated test systems (Figure 1) are supported by a DEC PDP 11/35 computer which utilizes a timeshare operating system. A valuable feature of this operating system utilized in this application,is the ability to access data being stored by executing an independent data analysis program at another remote terminal,without interrupting the data collection process.

The frequency stability certification test (Figure 4) is performed using test stations one and three (Figures 2 and 3). This test requires that data be collected for nine days without interruption. Microcomputers are included in these test stations to provide redundant data collection and storage if the main computer should fail. At the end of the test, these data would then be transferred to the PDP 11/35 for analysis. Failsafe devices are incorporated into all test stations to protect against RFS supply over-voltage and current, and over and under base-plate temperatures.

The RFS supply current and telemetry-monitored testpoint levels are recorded throughout the test cycle. The RFS frequency is measured using a "Time Mark System. [2] This system, developed by ESG Metrology, is similar in principle to the NBS chronograph. The heterodyned beat period is measured without dead time (the loss of time required to rearm the counter). In this system, the beat-frequency zero crossings load the value of a free running counter into an output buffer. This system was implemented by simple modification of an electronic counter.

Figure 5 shows the RFS mounted in the vacuum chamber which similates the spacecraft pressure. Cooling is provided through the RFS baseplate to the mounting block which simulates the spacecraft mounting. Coolant from a temperature-stabilized bath is circulated through a maze within the block.

The RFS is fundamentally an Efratom Rubidium Frequency Standard repackaged with extensive modification to meet space environmental, reliability and spacecraft interface requirements. The RFS is 8 in. x 5 in. x 4.5 in. in size, weights 8-1/2 lb and consumes less than 30 watts of power. The nominal output frequency is 10.23 MHz.

RFS PERFORMANCE

The data collected to obtain the two-sample Allan Variance [3,4] computer printout of Section 1, Figure 6, is for a 1-second fundamental beat period. Adjacent data pairs are averaged to obtain the 2-second Allan Variance, and so on for the 4- and 8-second results. The second section fundamental beat is 10 seconds. The third and fourth section results are calculated from stored data that is the average of 10 (100-seconds) and 80 (800 seconds) 10-second beat periods. This method is a reasonable alternative to storing large amounts of data, and still retaining acceptable confidence limits. The calculated drift per day is based on the last 5 days of testing. The Allan Variance results presented in Figure 6 reflect the removal of a first-order curve (drift).

Figure 7 is a graphical representation of the 800-second data used to calculate the Allan Variance of Figure 6. The graphs are the average of 10, 20, and 50, 800-second data, resulting in 8000, 10,000 and 40,000 seconds per bar. The graph readily demonstrates the exponential warmup characteristic present in the RFS frequency.

Figure 8 is a graphical representation of the printout of Figure 6. The vertical bars at each data point represent the range of confidence limits[3,4]. The inset in the top-right corner is the graph of Figure 7 for the average of 50, 800-second data points. It is apparent, that the warmup characteristic displayed by the inset is the predominant cause of the Allan Variance flattening or turning upward for longer sample times. In Figure 9 where the inset shows essentially flat data, or in Figure 10 where removal of the drift results in a relatively flat plot, the Allan Variance values continue a downward trend.

Figure 12 is a plot of the RFS frequency versus mounting base-plate temperature data. This test simulates possible RFS base-plate temperature excursions during a 12-hour space vehicle orbit. The accumulated time errors and the elapsed times over which they are specified are marked on the plot. The computer printout of Figure 13 shows the time error accumulations and the procurement specifications. Representative profile plots of other units are shown in Figures 14 and 15.

Figure 16 is a plot of the Allan Variance calculated from data collected in a RFS life cycle test conducted by the Space Systems Group of Rockwell International. The calculations reflect removal of the first-order drift characteristic. The test term was 100 days (8600 data of approximately 1000-second averaging time). A graph of the data revealed two distinct drift characteristics separated near the midpoint of the test. Both are very nearly linear. It seems likely that this change in the drift slope was caused by an

external influence on the RFS. The plot using the · symbol is the Allan Variance calculated from data points 1-8600, and the plot using the x symbol is calculated from data points 4300 - 8600.

These data indicate that the RFS, after adequate stabilization, is capable of meeting the GPS Phase II/III specification as shown in the figure.

ACKNOWLEDGEMENTS

The authors wish to express their sincere appreciation to Laurie Baker of Metrology, who made significant contributions in the development of the vacuum test systems and the temperature-control system for the Rubidium Frequency Standards.

REFERENCES

1. F. K. Koide and E. M. Hicks, "Atomic Clock Test Program for NAVSTAR," Fifth Cal-Poly Measurement Science Conference, December 1975.

2. E. M. Hicks, "An Innovative Method for Measuring Frequency Stability without Deadtime," Sixth Cal-Poly Measurement Science Conference, December 1976.

3. J. A. Barnes et al, "Characterization of Frequency Stability," NBS Technical Note 394, Issued October 1970; IEEE trans. on Instr. & Meas., IM-20, 1971.

4. P. Lesage and C. Audoin, "Characterization of Frequency Stability," IEEE Transactions on Instr. & Meas., IM-22, June 1975.

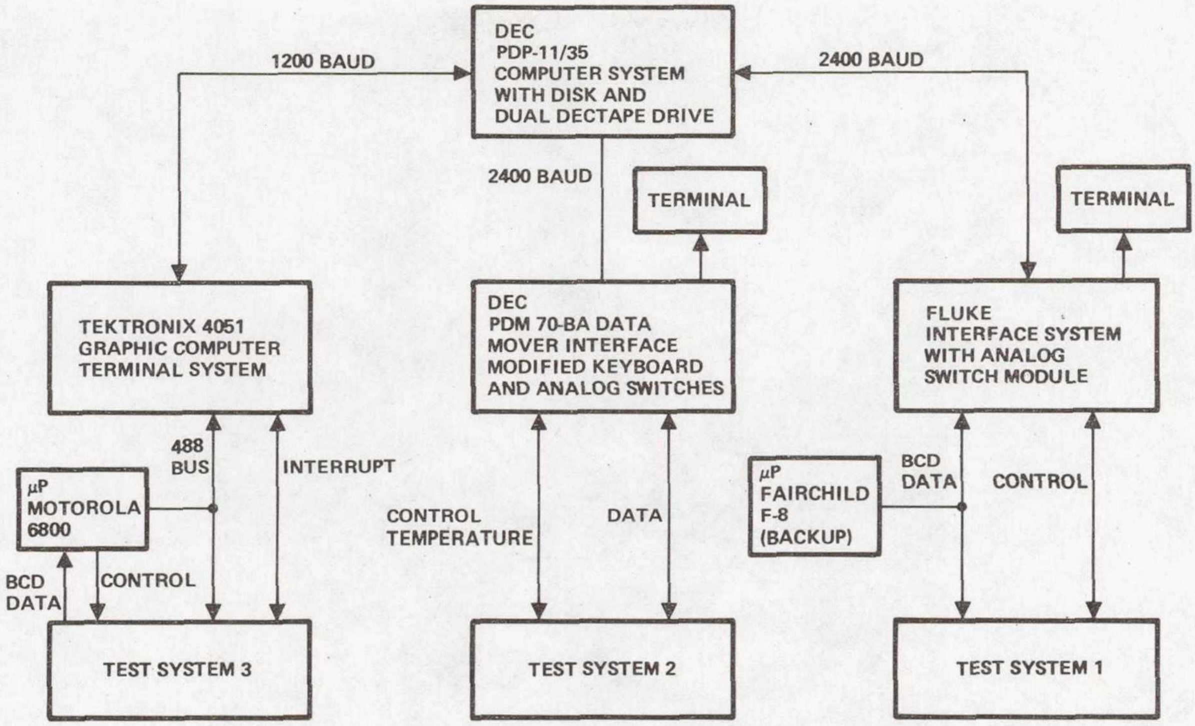


Fig. 1—Overall Automated Test Systems

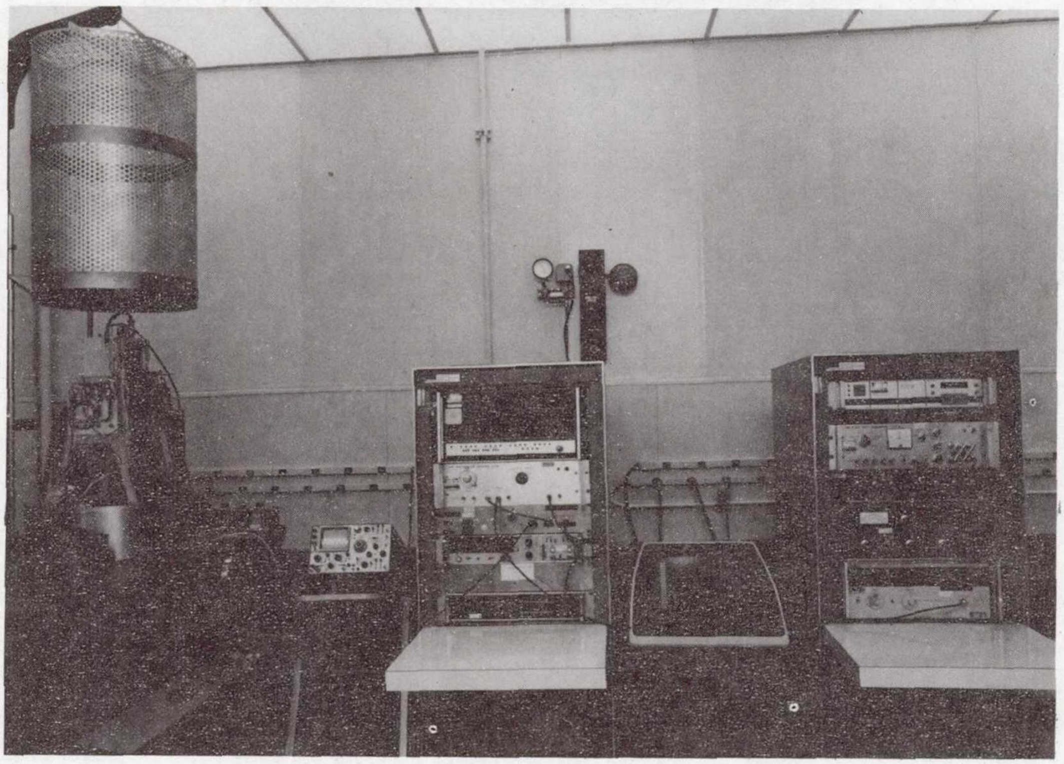


Fig. 2—Test Systems 1

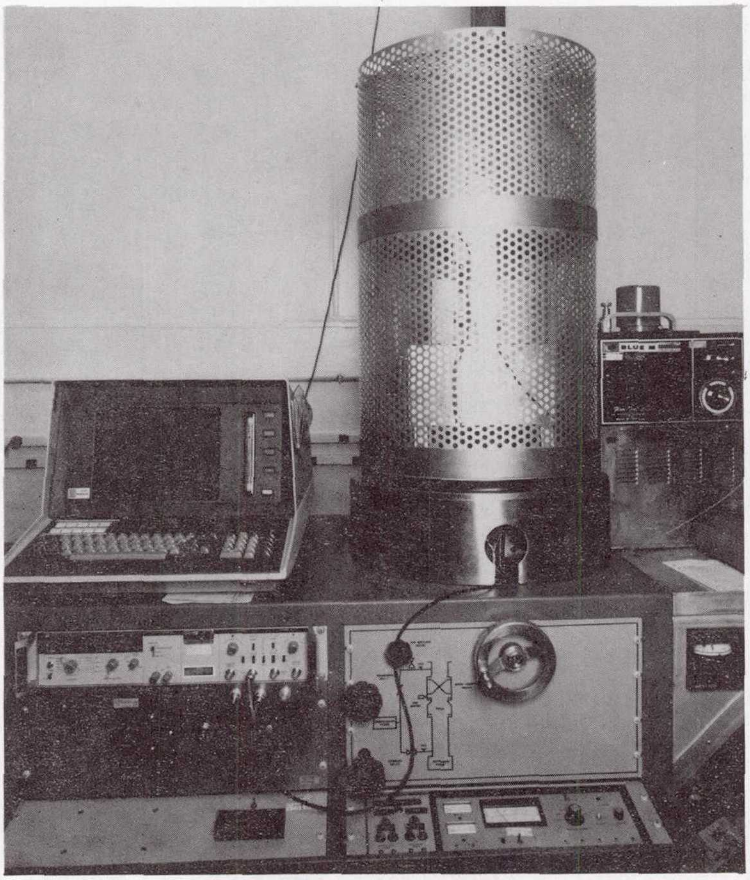


Fig. 3—Test System 3

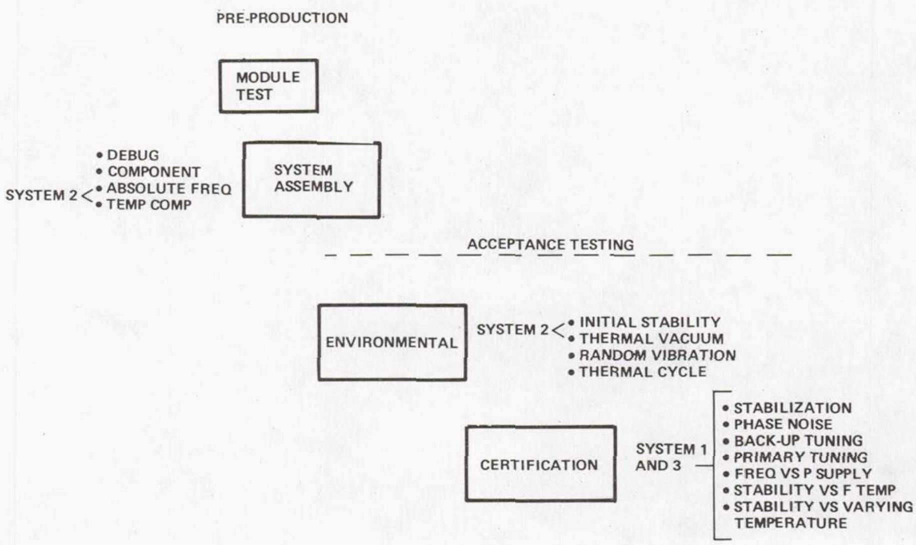


Fig. 4—Testing Cycle

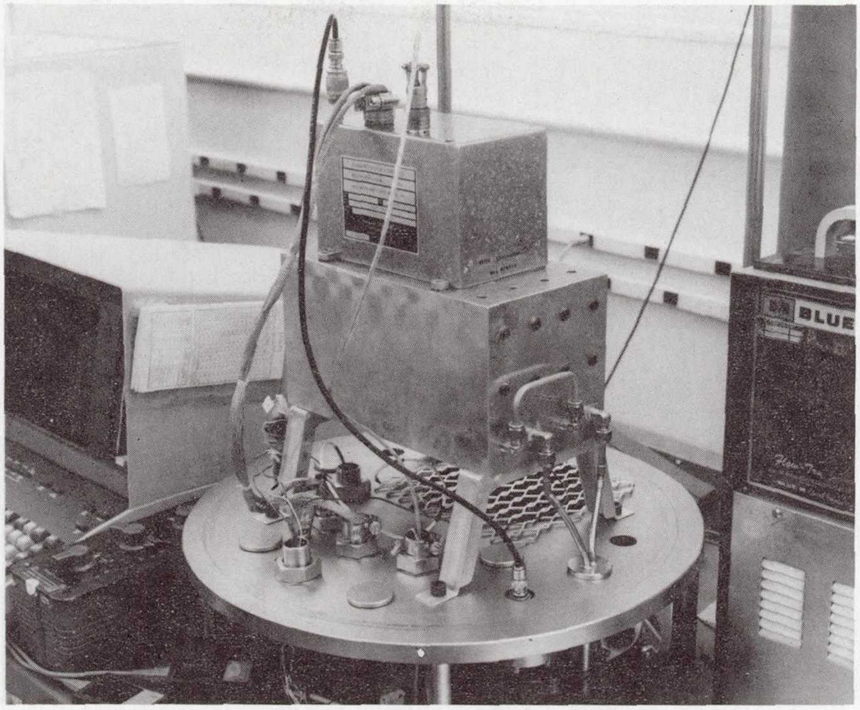


Fig. 5—Rubidium Frequency Standard

```
RUN
NEAL    15:37             31-OCT-78
SERIAL NO. OF RFS? 006
CURRENT COUNT ? 1200


LONG TERM STABILITY TEST  07-OCT-77  08:33

TEMPERATURE HELD BETWEEN  34.7  AND  35  DEGREES CELSIUS


        ALLAN VARIANCE CALCULATIONS
```

F	SAMPLES	TAU	SIGMA	CONF LOW LIM	CONF UP LIM
1	100	1.00	6.39E-12	5.74E-12	7.03E-12
2	50	2.00	6.39E-12	5.48E-12	7.30E-12
4	25	4.01	5.92E-12	4.71E-12	7.13E-12
8	12	8.01	5.00E-12	3.49E-12	6.50E-12
1	100	10.37	3.25E-12	2.97E-12	3.53E-12
2	50	20.74	2.97E-12	2.60E-12	3.34E-12
4	25	41.48	2.14E-12	1.76E-12	2.52E-12
8	12	82.96	8.29E-13	6.12E-13	1.05E-12
1	100	103.68	1.25E-12	1.14E-12	1.35E-12
2	50	207.35	9.68E-13	8.48E-13	1.09E-12
4	25	414.70	6.16E-13	5.06E-13	7.25E-13
8	12	829.40	5.40E-13	3.99E-13	6.82E-13
1	900	828.26	4.42E-13	4.30E-13	4.55E-13
2	450	1656.51	3.06E-13	2.93E-13	3.18E-13
4	225	3313.02	2.14E-13	2.01E-13	2.26E-13
8	112	6626.05	1.49E-13	1.37E-13	1.62E-13
16	56	13252.10	1.50E-13	1.26E-13	1.74E-13
32	28	26504.19	1.60E-13	1.23E-13	1.97E-13
64	14	53008.38	2.24E-13	9.32E-14	3.55E-13
128	7	106016.77	3.67E-13	8.26E-14	6.52E-13
256	3	212033.54	5.88E-13	0.00E 00	1.19E-12

```
REMOVES DRIFT

DRIFT/DAY=      .377746E-12

READY
```

Fig. 6—Computer Printout of Allan Variance Calculation

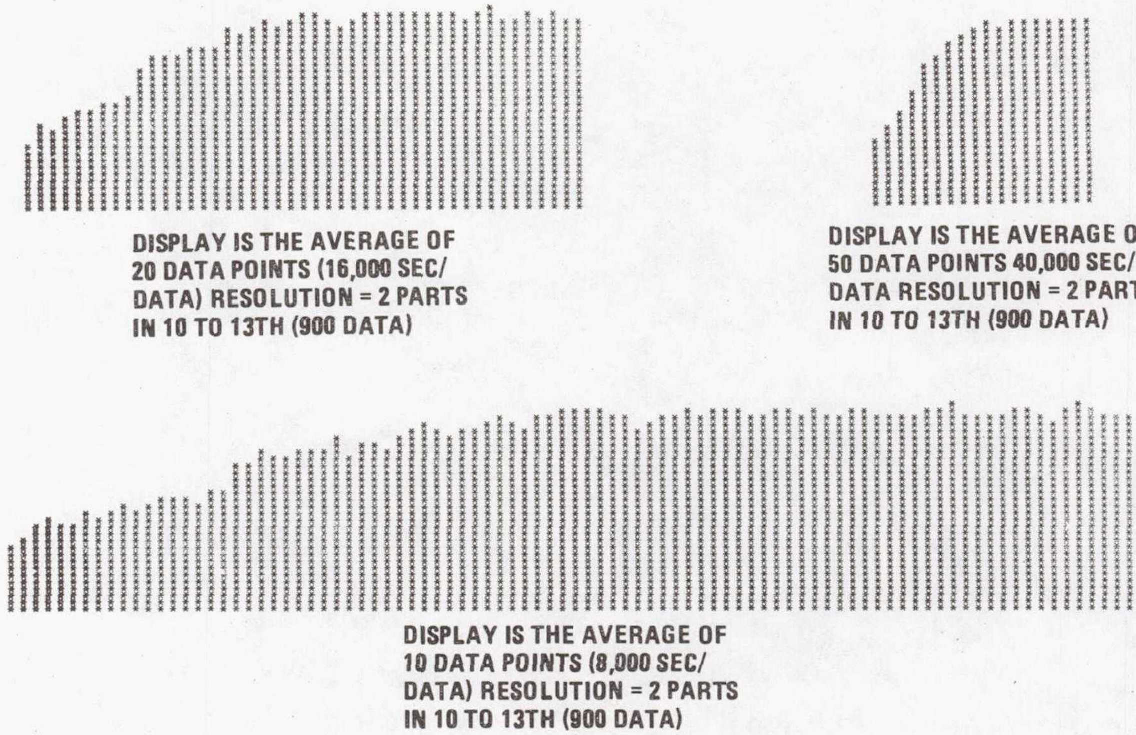


Fig. 7—Graphical Plots (SN 6)

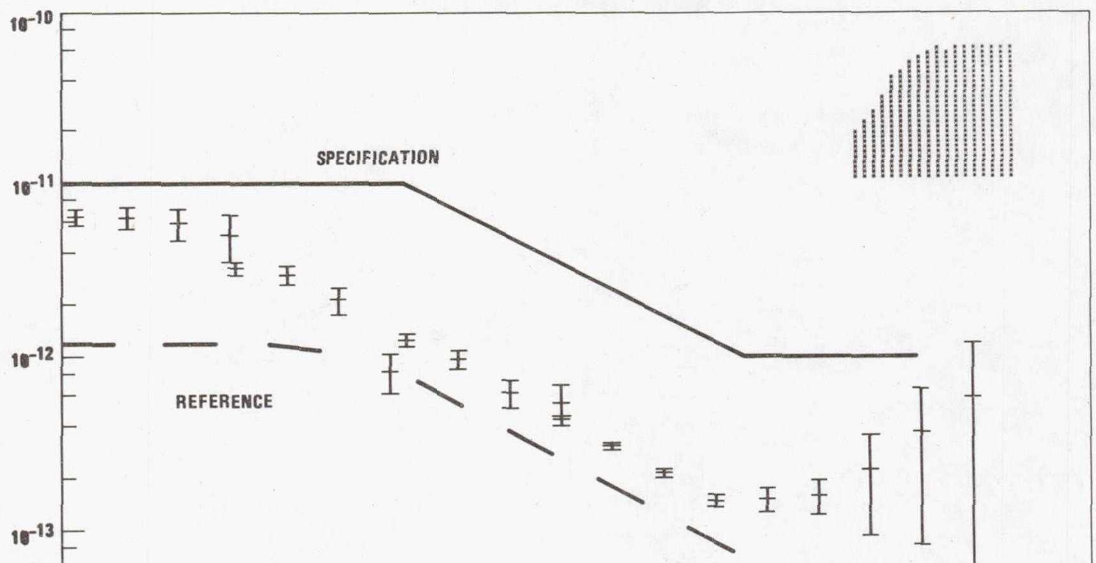


Fig. 8—Computer Printout of Allan Variance Plot

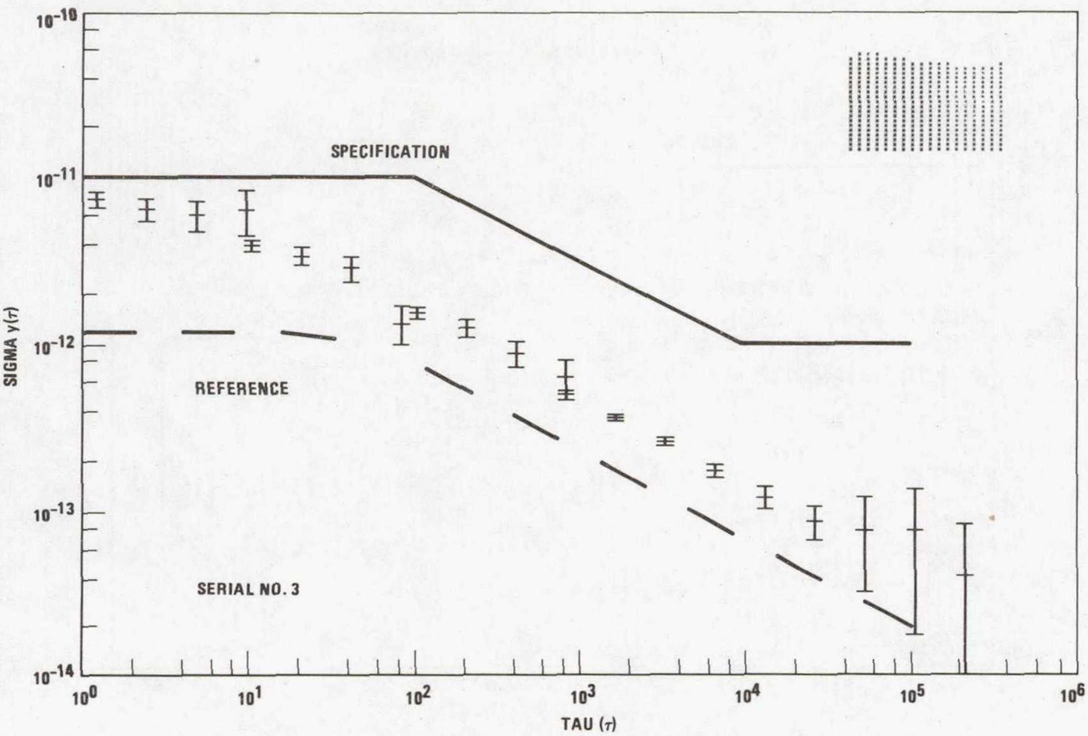


Fig. 9—Computer Printout of Allan Variance Plot

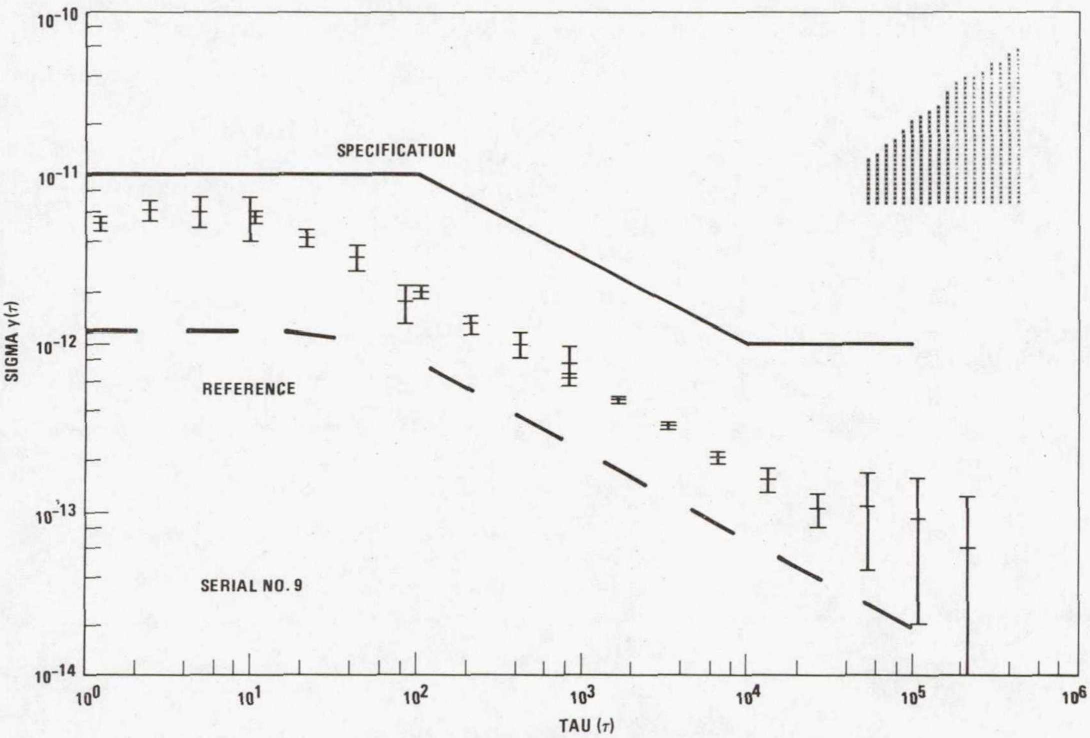


Fig. 10—Computer Printout of Allan Variance Plot

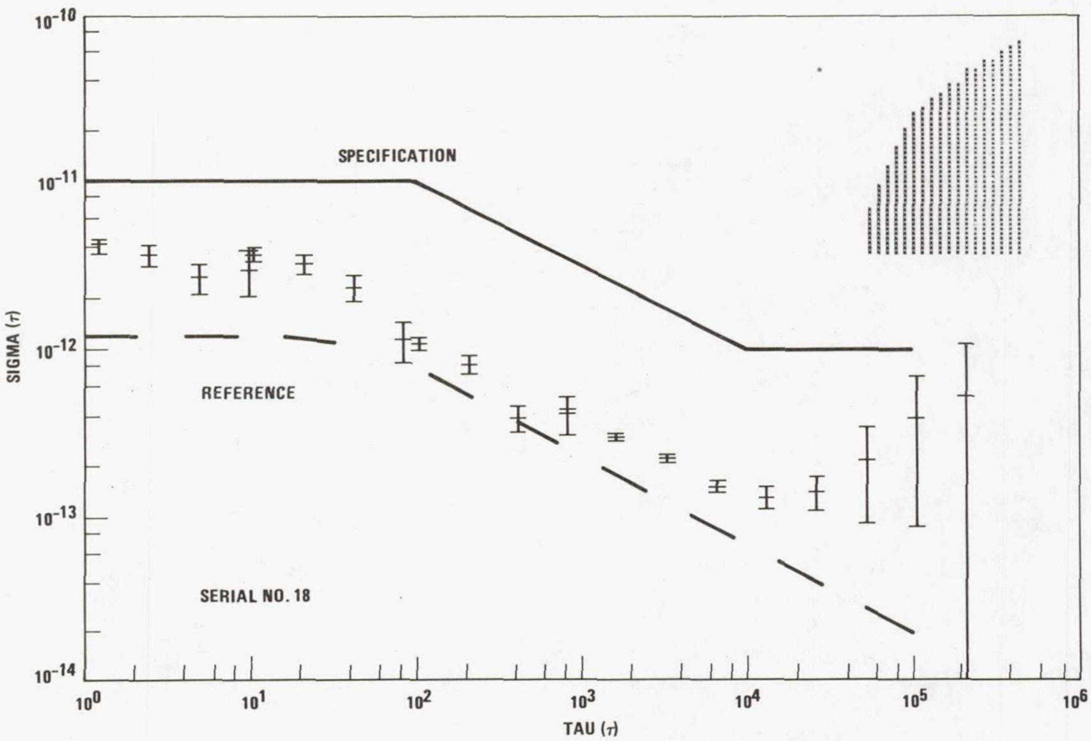


Fig. 11—Computer Printout of Variance Plot

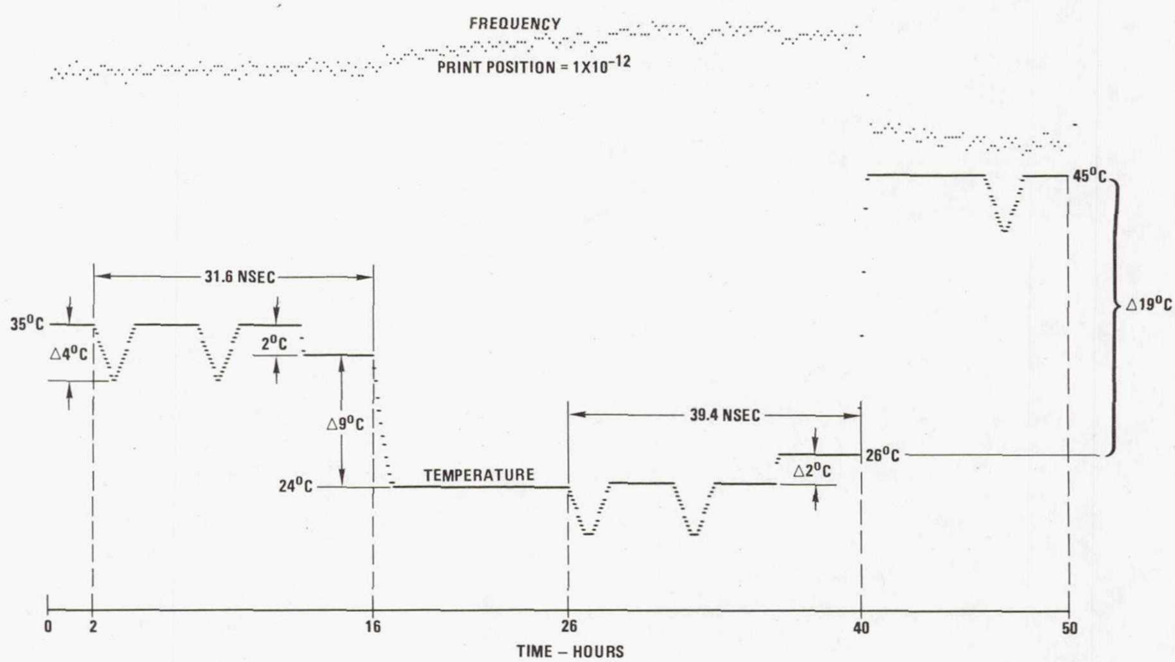


Fig. 12—Computer Plot

RUN
F 10:17 06-NOV-78
SERIAL NO OF RFS? 003

TEMP/STABILITY TEST 29-JUN-78 14:16

ELSPD TIME	DEG-C TEMP	N-SECS PHASE	L-LIMIT PHASE	U-LIMIT PHASE
0.00	35.1	0.00	-88.00	88.00
0.43	35.1	0.21		
0.86	34.8	0.61		
1.30	33.0	1.65		
1.73	31.3	2.35		
2.16	32.4	2.41		
2.59	34.1	1.68		
3.03	35.1	2.61		
3.46	35.1	2.35		
3.89	35.1	3.14		
4.32	35.1	4.46		
4.76	35.1	5.29		
5.19	35.1	6.03		
5.62	35.1	6.76		
6.05	34.4	8.46		
6.49	32.7	8.87		
6.92	31.1	10.12		
7.35	32.8	10.70		
7.78	34.5	11.34		
8.22	35.1	11.71		
8.65	35.1	13.66		
9.08	35.1	14.53		
9.51	35.1	15.42		
9.95	35.1	17.20		
10.38	35.1	18.87		
10.81	35.1	19.69		
11.24	33.0	21.92		
11.68	33.0	23.68		
12.11	33.0	25.69		
12.54	33.0	26.90		
12.97	33.0	28.31		
13.41	33.0	30.28		
13.58	33.0	30.97		
13.67	33.0	31.11		
13.75	33.0	31.22		
13.84	33.0	31.31		
13.93	33.0	31.37		
14.01	33.0	31.64		

ELSPD TIME	DEG-C TEMP	N-SECS PHASE	L-LIMIT PHASE	U-LIMIT PHASE
25.00	23.7	0.00	-75.00	75.00
25.43	23.6	0.40		
25.86	22.4	0.51		
26.30	20.7	1.94		
26.73	20.0	-0.17		
27.16	21.9	-2.09		
27.59	23.6	-1.98		
28.03	23.7	-0.92		
28.46	23.7	0.68		
28.89	23.7	2.21		
29.32	23.7	3.74		
29.75	23.7	6.54		
30.19	23.7	8.81		
30.62	23.7	11.05		
31.05	22.1	13.80		
31.48	20.4	15.33		
31.92	20.3	14.55		
32.35	22.3	14.68		
32.78	23.7	16.84		
33.21	23.7	19.52		
33.64	23.7	21.76		
34.08	23.7	24.37		
34.51	23.7	26.92		
34.94	23.7	29.87		
35.37	23.7	32.85		
35.81	25.6	35.34		
36.24	25.7	34.87		
36.67	25.8	34.24		
37.10	25.8	34.98		
37.53	25.8	35.97		
37.97	25.8	37.47		
38.40	25.8	38.21		
38.83	25.8	39.56		
39.00	25.8	39.66		
39.09	25.8	39.39		

a b

Fig. 13—Computer Printout of Frequency vs Temperature Profile Test

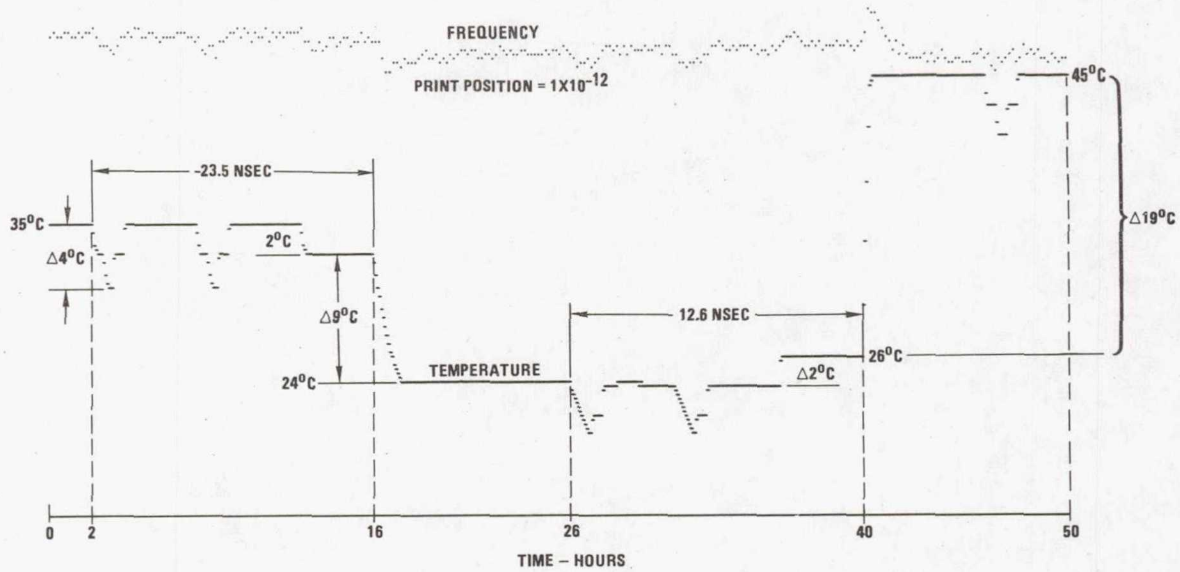


Fig. 14—Computer Plot

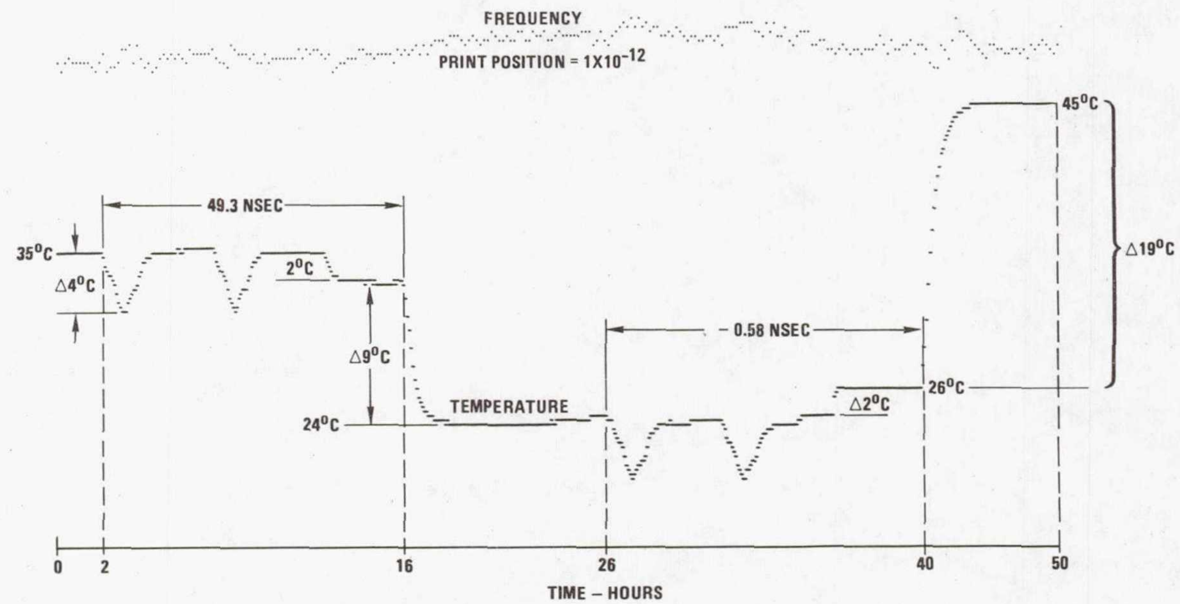


Fig. 15—Computer Plot

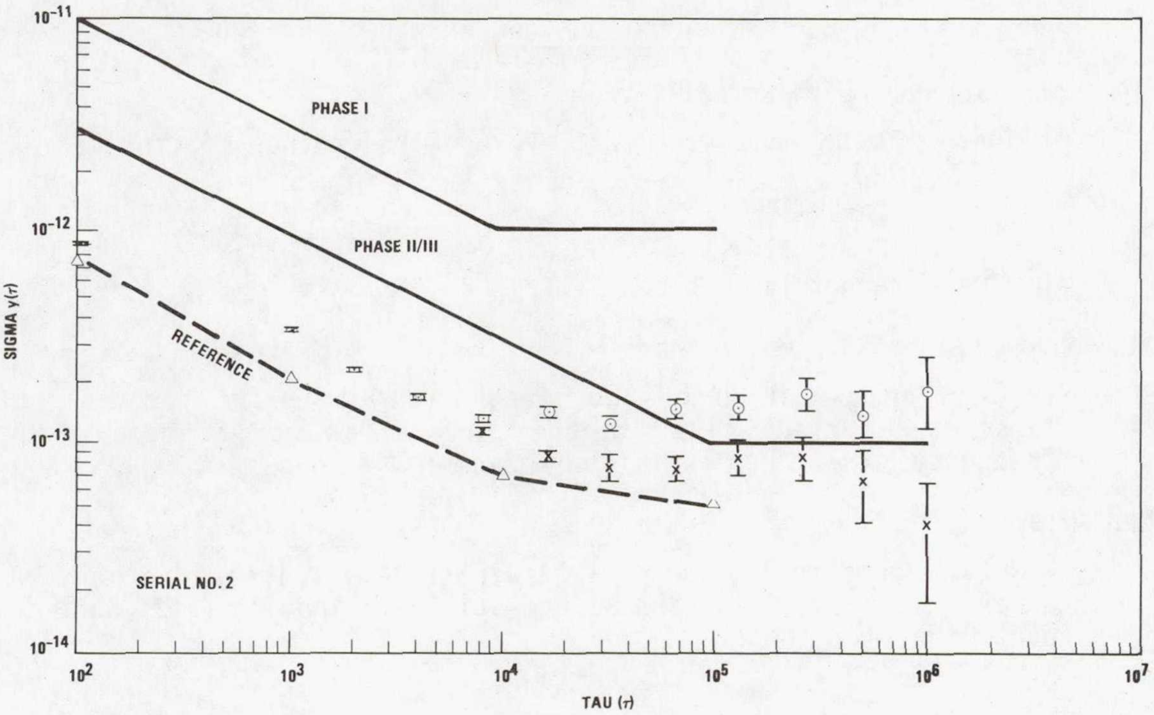


Fig. 16—Allan Variance Plot/Life Cycle Test

QUESTIONS AND ANSWERS

MR. HUGO FRAUHOFF, Efratom California:

I just want to make sure that everyone knows that Efratom is teamed with Rockwell.

MR. KOIDE:

I think I forgot to mention that.

MR. ANDREW CHI, NASA Goddard Space Flight Center:

In your presentation of the frequency plot versus temperature or temperature versus time, in one curve there was a sudden jump in frequency. Could you explain why there was a large jump?

MR. KOIDE:

Yes, we don't really have an answer to that. There are many variables that cause that, and we are still looking into it. So I don't want to go into that.

OVERVIEW OF TIMING/SYNCHRONIZATION FOR DIGITAL COMMUNICATIONS SYSTEMS

Harris A. Stover
Defense Communications Engineering Center, Reston, Virginia

ABSTRACT

Digital technology is being applied to communications at a very rapidly increasing rate. Use of digital communications (transmission, multiplexing, and switching) results in a timing/synchronization problem not encountered in analog communications systems. This overview explains the need for timing/synchronization of digital communications systems in general, and switched systems in particular. It points out some of the criteria that greatly influence timing/synchronization subsystem design for a military communications network but have little or no significance for civil systems.

The results of previous studies by DCA and its contractors are summarized and the rationale for the current study being reported in other presentations in this session is provided. In this approach, timing techniques are evaluated in terms of fundamental features. Different combinations of these features would cover most possibilities from which a synchronous timing system could be chosen. Although the studies described in this session were intended for application to the Defense Communications System, the problems of an extensive worldwide military communications system tend to encompass the problems of other digital communications systems, both military and civilian.

INTRODUCTION

The communication of information in digital form is expanding at a very rapid rate. There are many reasons for this: (1) to accommodate information that originates in digital form; (2) economic advantages of digital multiplexing and switching equipment; (3) improved quality provided by digital telephone systems (noise and distortion do not accumulate along the path); (4) need for communication to, from, and between digital computers; (5) requirement for encryption of communications; (6) available techniques to provide resistance to signal jamming; (7) reduced maintenance costs for digital systems, and (8) the trend whereby developing technology continues to further enhance the advantages of digital communications. There are timing/synchronization requirements that must be satisfied in order to reap the

advantages of digital communications. As will be discussed in the following sections, the timing/synchronization requirements encountered when digital communications are organized into networks can be satisfied by many different approaches. Information presented in the four papers of this session of the meeting should help to choose from among them.

DIGITAL COMMUNICATIONS

In digital communications, transmission is in the form of discrete pulses, each with a finite number of allowable states. For pulses with two states (e.g., two phases), each pulse represents a single bit of information; for four states (e.g., four possible phases), each pulse represents two bits; for eight states, each pulse represents three bits, etc. The number of bits is equal to the logarithm to the base 2 of the number of allowable states.

For a one-way digital point-to-point communications link, it is only necessary that the receiver be correctly synchronized to the received signal. However, since a different type of meaning can be applied to different pulses in the sequence--e.g., the most significant digit in a large number has a different meaning than does the least significant digit, or different pulses may represent data with different origins--a capability must be provided to identify particular pulses. This is normally done by grouping the bit streams into frames of a selected number of bits. Frame synchronization codes are used to identify the beginning of the frames. These frame synchronization codes are chosen so as to be unlikely to occur in the communications stream. This is accomplished by selection of unique patterns or by transmitting the code with greater regularity than it would randomly occur.

Most digital communications systems, other than simple telemetry systems or one-way command systems, provide transmission in both directions. By itself this does not impose much of a synchronization problem. Two-way transmission is effectively two one-way links in opposite directions. However, advantage can be taken of the two-way transmission to improve timing precision in more complex systems.

MULTIPLEXING

For economy, a number of lower capacity channels are multiplexed onto a single higher capacity transmission link. In digital transmission, this is done by interleaving the pulses from several lower rate channels into a single higher rate stream, i.e., time division multiplexing. In order to accomplish this without problems, the signal pulses must be available at the correct time to fill their assigned time slots in the multiplexed pulse stream, i.e., some form of synchronization must be provided.

PULSE STUFFING

One procedure for multiplexing digital communications streams, that may have clock rates that are slightly different, adds extra (dummy) pulses to each pulse stream to bring them all to a common higher rate compatible with the rate of the multiplexed transmission. At the other end of the transmission link, where the channels are demultiplexed, the extra pulses are removed to return each channel to its original rate. This technique, referred to as "pulse-stuffing," is quite effective for individual transmission links, but it does introduce some pulse jitter. In order to reduce the cost of pulse stuffing, it is common practice for telephone companies to group analog voice channels together in groups of 24 channels (30 channels in some countries) for digitizing at a common clock rate. When this is done, the pulse stuffing is only needed for time division multiplexing these groups of 24 channels into higher capacity transmission links. For this use, the cost of any single stuffing-destuffing operation is shared among all 24 channels. For more complex networks, where at any particular node previously multiplexed channels are demultiplexed, separated and remultiplexed with other channels on each of several links leading to different nodes, the requirements of pulse stuffing and destuffing are much more complex.

SWITCHING

If switches are provided in the network so that a channel originating at any user of the network can, upon demand (dialing and right number), be connected to any other user located anywhere in the network, the network is being reorganized nearly continually. The particular channels that are multiplexed together continually change. This means that every individual communications channel must be capable of being switched and/or multiplexed anywhere in the network on an individual channel basis. The telephone companies commonly provide this individual channel capability for transmission involving pulse-stuffing by returning all signals to analog form at all switches and redigitizing the signals after switching. This approach is relatively expensive. For encrypted digital voice signals, returning them to analog form would require considerable additional expense for encryption/decryption equipment at all switches. This decoding/encoding would also reduce the security provided by the encryption. An economical alternative, permitting the signals to remain in digital form throughout the network, is to synchronize the entire network and all of its users. That is, instead of only synchronizing the receiver to its received signal for each transmission link, the transmitters throughout the network are all synchronized with one another.

VARIABLE BUFFERS

In synchronizing the digital communications network, variations in

the signal transit times (time it takes for the signal to travel from one node to another) on the various transmission links must be accommodated. Even if the transmitters were perfectly synchronized, these variations could cause changes in the time that received signals would be available to fill their assigned time slots. The transit time variation can be accommodated by providing each receiver with a variable storage buffer (or stack) into which each bit is placed when it is received. It is removed by the local clock at the correct time to fill its assigned time slot. By enlarging these buffers over the size needed for the path length variations, they can also accommodate some error in the local clocks.

THE INDEPENDENT CLOCKS TECHNIQUE

If very stable clocks are used at each node, and variable buffers (or stacks) of sufficient size are employed, then the clocks can free run for a useful period of time before the buffers either overflow or empty. When this capability is coupled with provision for occasionally interrupting communications traffic to reset the variable buffers (or stacks), the timing/synchronization technique known as "independent clocks" results. This technique has been chosen for use at major nodes in U.S. tactical communications. It is also being used as a backup mode of operation in other systems, and it is generally recommended for backup use.

MASTER-SLAVE TECHNIQUE

Although pulse stuffing is widely used on an individual link basis, it is not normally used where the signals must remain in digital form throughout the network. In North America most commercial communications systems in which signals remain in digital form throughout the network use some type of master-slave system. This is also true of civil systems in many other parts of the world. In these systems, all nodes of the network are slaved either directly or indirectly to a master by phase-locking the local clock at each node to a selected received signal. It is a very obvious and straightforward technique.

MUTUAL SYNCHRONIZATION

A technique which has been widely studied, but has had little application to date, is called "mutual synchronization." In it, each node adjusts the frequency of its clock in such a way as to reduce the phase difference between itself and some weighted average of the phases of all signals received from its neighbors.

EXTERNAL TIME REFERENCES

If a time reference for synchronizing each nodal clock is obtained from a source external to the communications network, it is referred

to as the "external time references" technique. Several timing sources could be used. These include such systems as Loran-C, the NAVSTAR Global Positioning System (GPS), etc.

TIME REFERENCE DISTRIBUTION

In the time reference distribution technique, all nodes are kept within a specified time tolerance of the master node. Time reference information is transferred between all connected nodes with the effects of transmission time removed. There are three major functions that must be performed by the time reference distribution technique: (1) selection of the paths over which the time reference is distributed through the network, including selection of a new master when necessary; (2) measurement of the local clock's time error; and (3) correction of the time error in the local clock. The technique is generally considered to have the features (defined later in the paper) of directed control, double endedness, self-organization, and independence of the clock error measurement at one node from the clock error correction at any other node except for the master. Some of these features can also be used in some of the other techniques as will be discussed later.

COMBINATIONS OF TECHNIQUES

The above techniques can be used in various combinations. The independent clock technique is an excellent backup for other techniques. However, in the applications to U.S. military tactical switched digital communications, the master-slave technique is used as a backup to the independent clock technique. The master-slave technique can be very effective at lower levels of the timing hierarchy when some of the other techniques are used at the higher levels. Another combination synchronizes the lower levels of the multiplex hierarchy (all users are synchronized) while using pulse stuffing at the higher levels of the multiplex hierarchy. This permits the use of lower speed logic for the buffers and for some of the other timing system hardware. However, it can greatly increase the complexity of some timing techniques because of the very large increase in the number of signals that must be separately synchronized as compared with synchronizing at the highest levels of the multiplex hiearachy.

MILITARY VS CIVIL REQUIREMENTS

Several characteristics of a digital communications network deployed for military communications can greatly influence timing/synchronization system design for such a military application, but have little or no significance for civil systems. However, careful examination of designs most capable of accommodating these characteristics might also show the approaches to be desirable for civil application. Among these characteristics is the need for extensive application of encryption to

military communications. Other such characteristics relate to action taken by an enemy to intentionally disrupt our military communications. A military digital communications network, including its timing function, must be capable of surviving such attacks. Resistance to signal jamming must be provided. Systems using wideband spread spectrum techniques for this purpose, (with their very short pulses arranged in a pseudo-random manner) must be able to rapidly synchronize their receivers to the received signals, even in the presence of jamming. The time needed to acquire such synchronization can be greatly reduced through the use of precise timing to reduce the size of the search window. An extremely stable timing system that is relatively immune from perturbation would help to satisfy the above requirements.

In normal operation, failures of the timing function will occur very rarely and would not be expected to be geographically extensive. Therefore, in civilian networks capable of free running for a period of time following failures, any required reorganization of network timing needed to accommodate failures can be manually controlled or initiated from a central location. However, in a military network, because of the importance of the timing function and the possibility of simultaneous attacks on many parts of the network, the reorganization of the timing subsystem should be highly automated and distributed throughout the network. (A required centralized function becomes an attractive target for enemy action.) Similar characteristics apply to the monitoring and maintenance functions for military network timing capability since they also may be intentionally impeded by the enemy.

Briefly stated, a military network must be able to endure: (a) physical destruction of parts of the network, (b) enemy capture of part of the network, (c) equipment failures, (d) enemy spoofing and/or jamming, and (e) deliberate obstruction of maintenance and repair operations. Under these conditions it must be able to maintain acceptable operations (both communications and timing) including operations within portions of the network which become isolated from the rest of the network. It must be able to interoperate with other digital communications networks that might use different techniques. It must be monitorable, primarily to permit the early identification of timing problems long before they can interrupt communications. It must also be versatile to provide easy compliance with modification of plans (such as the introduction of new technology) and to provide convenient application to future problems. The endurance, interoperability, monitorability, and versatility must be provided economically with low life cycle costs.

These and other characteristics of a military digital communications network have been considered in studies of communications network timing for the Defense Communications System.

STUDIES BY DCA AND ITS CONTRACTORS

Initial studies of network timing at DCA investigated DCS requirements and tradeoffs between several alternative techniques for meeting the requirements [1]. These initial studies selected a mutual system called Discrete Control Correction as desirable for use in the DCS. In it, the amount of occupied buffer storage is monitored for each link received at a node. At periodic intervals a weighted average of the information thus obtained is applied as a control signal for the correction of the local clock. Additional studies of the Discrete Control Correction approach were made by Clarkson College under a contract from the Air Force. These studies, employing both analysis and simulation, showed that proper selection of the weighting coefficients would bring the clocks to a common average rate with satisfactory damping of perturbations and the resulting freedom from spontaneous oscillations. The study participants were enthused by their findings which were described in a number of papers [2,3,4,5,6].

While these studies were being conducted at Clarkson College, further studies at DCA indicated that distributing an accurate time reference through the communications network would have advantages not available to mutual systems. For military application, such an approach would be capable of self-reorganization following failures. In order to maximize its stability and minimize the propagation of errors through the network, it would have no closed loops. The effects of signal transit time (in comparing the time of clocks at neighboring nodes) would be removed to enhance accuracy and stability. A method for accomplishing this is called Time Reference Distribution [7]. The general concept of which was presented at this planning meeting in 1973 [8].

With two different concepts being considered for Timing/Synchronization of the DCS, the Clarkson College team was tasked to compare four synchronization techniques: Discrete Control Correction, Master Slave, Independent Clocks, and Time Reference Distribution. This study found that use of the Time Reference Distribution Technique at major nodes with slaving at minor nodes will best meet the DCS requirements. However, the study indicated that the use of independent clocks has considerable merit if it is acceptable to occasionally interrupt traffic to reset the variable storage buffers. But, the study also stated that the availability of monitoring capabilities renders the Time Reference Distribution Method superior to the Independent Clock Technique [9]. While this study was being conducted, work at DCA showed how easy it would be to make the measurement of the clock error at any node independent of the correction of the clock error at any other node by simply having each node inform its neighbors of its measured but uncorrected error [10, 11].

Because it was felt that an industrial contractor could add a degree of depth in the area of hardware implementation that could not be

provided as well by an educational institution, and because the Clarkson College report on evaluation modeling had recommended further study of some aspects of some of the timing techniques, the Harris Corporation was competitively selected for a "Study of Alternative Techniques for Communication Network Timing/Synchronization." The results of this study were in agreement with all previous studies that had considered the Time Reference Distribution Technique; i.e., that the Time Reference Distribution Technique should be used at the major nodes of the switched digital Defense Communications System, but that minor nodes of the network should be slaved [12]. There was criticism of the Harris Corporation report because it compared simple Master-Slave and Mutual systems with a Time Reference Distribution system that has many more features. The criticism pointed out that many of the features of the Time Reference Distribution Technique could be included in either the Master-Slave or Mutual Techniques. So long as too many of these features are not included, these systems will not actually be Time Reference Distribution Systems.

Simultaneously with this study, an Improved Time Reference Distribution Concept was being developed at DCA. This improved concept provides automatic selection of the highest ranking clock in the network as the master for the network; automatically arranges the network into a preferred timing hierarchy below that master; provides optimum combining of the timing information at each node for best accuracy while avoiding all closed paths; and provides independence of the clock error measurement at any node from a clock correction at any other node. This independence prevents changes or adjustment of any clock in the network--other than the master clock--from propagating to other nodes [13]. This is accomplished without requiring any node to communicate with any other node farther away than its immediate neighbors.

It had been planned that studies following the "Study of Alternative Techniques for Communications Network Timing/Synchronization" would optimize a selected system. However, as a result of the criticism of that study, it was decided that the follow-on study should include further comparisons. In order to avoid further criticism, the Statement of Work for the follow-on "DCS Synchronizing Subsystem Optimization/Comparison Study" was written to require a comparison of the capability of each of a number of basic subsystem features to provide each of a number of desirable characteristics. This would provide a basis of comparison for any combination of these features. These different combinations would cover nearly all meaningful possibilities for synchronization of digital communications systems. The features were to include as a minimum:

1. Directed Control: The clock at only one end of a transmission link is permitted to be changed as a result of the measured difference

between the clocks at the two ends.

2. Double Ended: information is exchanged between the two ends of a transmission link so that the clocks at the two ends can be directly compared independent of the time required for the signal to travel from one node to the other. (If the time required for the signal to travel between nodes is the same for both directions of transmission, this feature permits removal of the signal transit time from the comparison of the two clocks).

3. Self Organization: when the network is initially put into operation, or following a disturbance, the synchronization function of the network automatically organizes itself into an optimum configuration for all surviving facilities.

4. Independence of the clock error measurement at any node from the clock error correction at any other node: all clock error measurements are, in effect, made with reference to the master; and errors in any clock other than the master do not propagate through the network.

5. Overhead: while not a feature as such, all systems (except for very low rate systems not considered here) regardless of features have timing subsystem overhead requirements for providing the frame synchronization signals (even the independent clocks technique). How much this basic overhead requirement is either increased or decreased by implementing the other features is of interest.

The desirable characteristics against which the features were to be evaluated included as a minimum:

1. Survivability of the Timing Function.

2. Slip-free operation, i.e., no interruptions of traffic to reset the variable storage buffers.

3. Frequency accuracy and phase accuracy.

4. Freedom of clocks from disturbance by perturbations at clocks or transmission facilities further from the master than the clock under consideration.

5. No harmful propagation to any other node of an error introduced at any node in the network except the master.

6. Compliance with Federal Standard 1002, "Time and Frequency Reference Information in Telecommunications Systems." (This requires referencing to UTC (USNO) or UTC (NBS).)

7. Monitorability at the system level (functional vs equipment

monitoring).

8. Minimum overhead communications.

9. Interoperation of the digital communications system with other digital communications systems employing different synchronizing techniques.

10. Cost effectiveness, i.e., maximizing the ratio of performance to cost.

11. Capability of automatic selection of a new master whenever there is a failure of the master through which the timing/synchronization subsystem is coordinated.

12. Availability of Precise Time (UTC) to users of the DCS without introducing any significant penalty.

The results of this work (2 tasks of the 11 task study) are reported in the third and fourth papers of this session.

EXPERIENCE WITH OPERATIONAL SYSTEMS

It is recognized that frequently much can be learned during the initial stages of implementing a system, and that this frequently includes valuable information that does not come to light during theoretical studies only involving analysis and simulation. Because of this, one task of current study required gathering information on the implementation and operational experience with timing/synchronization systems of operating digital communications networks. The results of this task are reported in the second paper of this session.

REFERENCES

[1] DCASEF Report 720.5-2 "Network Timing and Synchronization," R. H. Bittel, 17 May 1972.

[2] R. Mukundan et al, "Clock Synchronization Through Discrete Control Correction," EUROCON '74 Conf. Digest, Amsterdam, The Netherlands, April 1974.

[3] K. R. Krishnan et al, "Synchronization of a Digital Communication Network by Discrete-Control-Correction," IEEE IASACS '74 Proceedings, April 1974, pp 411-415.

[4] D. A. Perreault et al, "A Computer Simulator for Discrete Control Correction," Fifth Annual Pittsburg Conference on Modeling and Simulation Proceedings, Pittsburg, PA, April 1974.

[5] R. H. Bittel et al, "Clock Synchronization Through Discrete Control Correction," IEEE Trans. on Communications, 22, No. 6, June 1974, pp 836-839.

[6] K. R. Krishnan et al, "Discrete Control Correction for Synchronization of Digital Communication Networks with Different Correction Parameters at Different Nodes," IEEE ICC '74 Conf. Record, June 1974, pp 6C-1--6C-4.

[7] DCASEF TC 39-73, "Time Reference Concept for the Timing and Synchronization of the Digital DCS," H. A. Stover, July 1973.

[8] H. A. Stover, "A Time Reference Distribution Concept For a Time Division Communications Network," Proceedings of Fifth Annual NASA Department of Defense Precise Time and Time Interval (PTTI) Planning Meeting, NASA, Goddard Space Flight Center, Greenbelt, Maryland, December 4-6, 1973.

[9] RADC-TR-77-19 (AD-039-377) "Network Timing/Synchronization Evaluation Modeling," Clarkson College of Technology, April 1977.

[10] H. A. Stover, "Coordinated Universal Time (UTC) as a Timing Basis for a Digital Communications Network," EASCON 1974.

[11] DCEC TR 43-75 (AD-A021934) "Communications Network Timing," H. A. Stover, September 1975.

[12] Harris Corporation Electronic Systems Division, AD-A041004 "Final Report, Study of Alternative Techniques for Communication Network Timing/Synchronization," March 23, 1977.

[13] H. A. Stover, "Improved Time Reference Distribution for a Synchronous Digital Communications Network," Proceedings of Eighth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, U.S. Naval Research Laboratory, Washington, D.C., November 30 - December 2, 1976.

Page intentionally left blank

SURVEY OF TIMING/SYNCHRONIZATION OF OPERATING WIDEBAND DIGITAL COMMUNICATIONS NETWORKS

RONALD L. MITCHELL*
NEC AMERICAN, INC.
7515 GREENVILLE AVENUE, SUITE 400
DALLAS, TEXAS 75231

ABSTRACT

In order to benefit from experience gained from the synchronization of operational wideband digital networks, a survey was made of three such systems: Data Transmission Company, Western Union Telegraph Company, and the Computer Communications Group of the Trans-Canada Telephone System. Additional information was obtained from AT&T relative to their Switched Digital Network.

The focus of the survey was on deployment and operational experience from a practical (as opposed to theoretical) viewpoint. The objective was to provide a report on the results of deployment-how the systems performed and wherein the performance differed from that predicted or intended in the design. It also attempted to determine how the various system designers would use the benefit of hindsight if they could design those same systems today.

No conclusions or evaluations of the network performance are provided in the report. However, some of the differences in requirements between the commercial networks surveyed and those of strategic, survivable networks, such as the Defense Communications System, are noted.

INTRODUCTION

There are (were) several sideband digital communications networks operating which incorporate synchronization systems. Most of these networks were originally conceived and developed in the late sixties and early seventies, and initially deployed in the early seventies. This paper describes the results of a survey which attempted to learn of practical aspects and experiences in the design, deployment and operation of these networks.

Three specific systems surveyed are the switched digital network developed by the Data Transmission Company and now operated by SP Communications (hereafter called the Datran System); the digital network deployed and operated by the Computer Communications Group of the Trans-Canada Telephone System (Dataroute) and a digital system developed by the Western Union Telegraph Co. (hereafter called the WU system). With the

*Dr. Mitchell was with Electrospace Systems, Inc. when the study was conducted under a subcontract from the Harris Corporation.

exception of the WU System, these networks are in service today. The Western Union system was put into operation in late 1971 and field trials conducted for several months. The system was never put into service, and has since been dismantled.

A brief description of each major subsystem of the three networks is given. It is followed by a set of questions which were posed to personnel associated with each of the networks in the development and early operational phases. These questions were posed during meetings with the personnel for each of the systems. Original plans were to include a section on the private-line Digital Data System (DDS) deployed by American Telephone and Telegraph. However, due to pending litigation, AT&T personnel were unable to provide information on DDS. An interview was substituted which covered Bell's planning for timing and synchronization of the network of No. 4ESS switching centers referred to as the Switched Data Network (SDN).

The paper concludes with a summary of the survey and a general comparison of timing/synchronization (T/S) considerations for the commercial networks described with those for a strategic military network such as the DCS. Many of the features important to a military system such as the DCS need not be considered in a commercial system. Survivability may be of interest to a commercial network, but not enough to dominate design trade-offs. Thus, while the engineering, planning, and design of the commercial networks, along with the operational experiences, are of interest in planning military networks, the two applications have distinctions that can heavily sway design trade-offs and decisions in different ways.

Questions posed to the technical personnel associated with the three digital systems (WU, Datran, Dataroute) were as follows:

A. Fundamental Reasons for Choosing Timing and Synchronization (T/S) As Implemented.
   1. What were criteria used for selecting T/S system?
   2. Were other types T/S system considered?
   3. On what basis was final choice made-performance analysis; cost analysis; O/M analysis; simulation; other?
   4. Is system reference to UTC or other standard necessary and/or desirable?
   5. How was compatibility with other systems (international, military, other common carriers, etc.) a factor in the system selection/design?

B. Theoretical Design of Chosen T/S System
   1. What were the main criteria in specifying the T/S system? Did it have to "marry" existing hardware, etc.?
   2. Specifications of Hardware

(a) Stability - long/short term
(b) Reliability
(c) Jitter reduction
(d) Cost goals
(e) O/M aspects

3. What were driving criteria on timing distribution waveforms, interfaces, etc. "downstream" from clocks, i.e., between MUX's, from MUX's to user, etc.?

C. Practical Design/Engineering Considerations
1. Were any unusual obstacles met in the design/development of hardware?
2. Were specifications as finally met tighter or looser than original specifications?

D. Initial Field Deployment, Testing, O&M
1. In the field trials and early operation of the network, were any unusual problems encountered?
   (a) Were craftsmen able to install, test, troubleshoot, or was system so foreign that engineers were needed?
   (b) Was elaborate/non-routine test equipment needed?
2. Are there any changes to system design that, in retrospect, would enhance or otherwise favorably influence initial field deployment and testing?

E. System Operations Experience
1. Has system operated as expected? If not, what has arisen that was unexpected?
2. Is performance adequate in retrospect?
3. Is reliability adequate in retrospect? Is redundancy used?
4. Is maintainability adequate in retrospect? What have been MTTR experiences?
5. Is flexibility for growth/change acceptable in retrospect? Are any future requirements <u>now</u> foreseen that would have influenced design?

F. Any Other Information Useful to U.S. Government/DoD in Planning for the Defense Communications System (DCS)?

<u>WESTERN UNION DIGITAL NETWORK</u>

The Western Union Telegraph Company network in this survey was operated as a testbed until it was dismantled in late 1973 or early 1974. The system utilized line-of-sight microwave as the basic trunking medium and provided point-to-point data services. The system, while only deployed over a limited area (Cincinnati to Atlanta) was conceived and designed to be eventually expanded on a nationwide basis.

The system utilized partial response digital signals in the lower

baseband portion of an existing analog microwave radio system. Simultaneously, 600 FDM voice channels occupied the upper baseband.

The system was originally planned and designed around a five-level time division multiplex hierarchy. The fifth level of multiplexing (20 or 40 Mb/s output) was never implemented.

Both the planned and implemented systems employed a hybrid synchronizing arrangement wherein the first (or lower) two levels of multiplexing were synchronous system-wide while the top two (or three) levels of multiplexing employed pulse stuffing. This design, using asynchronous operation for the higher level multiplexers, minimized the use of high speed elastic stores, or buffers. Path length changes and other multi-bit delay variations were taken care of by the pulse-stuffing capability. Thus, the network was a hybrid, i.e., it was not purely synchronous or asynchronous.

The timing and synchronization system likewise was a hybrid in that it was conceived as a mixture of both master-slave and independent master types of systems. The segment of the planned system which was deployed between Atlanta and Cincinnati was basically operated as a master-slave system, but had the hybrid capability.

The timing and synchronization subsystem of the Western Union network was built around a set of redundant Disciplined Oscillators and a corresponding Interface Unit. The Disciplined Oscillators are basically phaselocked loops (PLL's) and the Interface Unit is a control circuit which selects the reference source to which the PLL's are locked.

For master-slave operation of the network, a highly stable oscillator such as a rubidium standard is fed into the Interface Unit at the master station. All other station clocks in the network then receive their reference from one or more of the 56 kb/s synchronous data channels emanating down from this master. At these slaved stations, the Interface Unit selects which 56 kb/s signal is to be used as the reference.

For independent master operation, each node in the system where timing is derived has a Loran-C receiver subsystem. This subsystem receives a ground wave signal from the Loran-C system and integrates it for a long period to provide a highly stable 1 MHZ signal to the Interface Unit and thence to the Disciplined Oscillators. All stations deriving timing from the Loran-C system are thus synchronized.

The Disciplined Oscillators are the heart of the Western Union timing and synchronization system. They were designed using a voltage-controlled crystal oscillator and a third-order feedback loop with

three time constants. The third time constant of approximately 200,000 seconds provides a highly stable output for several days in absence of an input reference. The third order loop remembers the rate of change of the difference in frequency between the internal VCXO and the reference frequency. When the input signal is removed, the feedback loop continues to correct for this difference. Thus, if the reference has remained unchanged, when it is reapplied, the drift rate of the VCXO will have been compensated as if the reference had not been lost. The other two time constants are 50,000 seconds for the second loop and (switch selectable) 5, 25, or 500 seconds for the first loop.

Because the Western Union system was never made fully operational, some of the questions asked were not directly answerable as noted below.

A. Fundamental Reasons for Choosing WU T/S System as Implemented

There were several criteria used in selecting the Western Union system as implemented. Cost was a primary consideration in choosing a synchronous system in that it was felt that the relatively large numbers of lower level multiplexers could be much simpler. The reason for choosing asynchronous higher level multiplexers was to avoid the relatively large costs of high speed elastic stores. A synchronous mode for the lower level multiplexers would also allow switched circuit operation, and this was the primary reason for selecting a synchronous network.

The master-slave technique, augmented by an independent master capability, was chosen for its simplicity of design and operation. At the time, it appeared that this technique, along with independent clocks, was well within the state of the art and achievable in practice, whereas systems such as mutual synchronization were considered (by Western Union) as unproven and questionable.

The independent clocks system was also considered, but it was concluded that the cost of avoiding periodically overflowing buffers was low enough to be well worthwhile. Programmed, or scheduled, error bursts associated with resetting buffers was judged unpalatable from a marketing point of view.

The final system choice was made on the basis of engineering judgment about practical aspects and associated technical risks. Implicit in this judgment was a performance, cost, and maintainability assessment. Simulations were not used.

Some brief consideration of using the network T/S system to distribute precise time was given. It was judged, however, that no market existed to justify the expense. Being able to synchronize the system to another system was judged desirable and was a factor in choosing the master-slave technique. It was felt at Western Union that eventually it would be desirable to synchronize with (slave to) other U.S.

common carriers.

B. Theorectical Design of WU T/S System

Several important aspects were considered when specifying the Western Union T/S system. It was initially planned to use a data signal-below-analog technique on existing analog radio systems and/or group band modems and/or leased T-1 facilities as the transmission media between nodes. Thus the system design, including the T/S equipment, had to be compatible with these transmission methods. The capability of the Interface Unit to accept 56 KHz, 1.544 MHz and 1 MHz reference signals reflects this requirement.

Another important criterion in specifying the T/S system was the required mean time to loss of bit count integrity at a node when the source of timing reference was lost. It was felt that bit count integrity should be maintained for up to several days if possible. To provide this, a stability of one part in 10^9 per day was chosen as the design goal.

Cost was also an important design criterion. In conjunction with the necessary reliability it was decided that the master-slave system augmented with a back-up Loran-C receiver/monitor was the most cost effective way to provide station timing throughout the network. An overall goal of Western Union was to develop the digital system to permit service at rates below the corresponding analog derived channels. The station clock system cost approximately $25,000 in 1972 dollars which was felt to be compatible with overall cost goals.

The following specifications were used for the development of the Western Union T/S equipment.

a. Stability - long/short term.
   Design goal of $\pm 1 \times 10^{-9}$ hours
b. Reliability - no reliability numbers such as MTBF were used. However, Western Union engineers felt that the T/S equipment should be of such reliability that overall system reliability would be determined by the radio system's MTBF and propogation path availability.
c. Jitter Reduction - no jitter reduction specifications were given.
d. Cost Goals - an approximate capital expenditure goal of $25,000 per station clock was used , not including spares provisioning or maintenance costs.
e. O/M Aspects - no MTTR or other maintenance related goals were specified.

The design of the signal/clock waveforms between the T/S system and the multiplexers and between the various multiplexer levels was

driven principally by the desire to be compatible with existing standards. In most cases, the standards were not formally specified in the early 70's but were de facto standards as perceived by the Western Union engineering staff. This basically led to using the T-1 bipolar format without separate clock signals wherever possible in the system.

C. WU Practical Design/Engineering Considerations

The only unusual or unforeseen obstacles encountered in the development of the Western Union system did not relate specifically to the T/S system. In the older radio system used between Atlanta and Cincinnati, some of the klystrons used as transmitter oscillators and as receiver local oscillators were microphonic. Mechanical vibrations were translated to electrical noise on the baseband in the lower frequency spectrum, which caused poor performance of the digital system. It was also found that phase and amplitude linearity of the radio system were critical, especially as the 600 channel voice system above the digital signal (on the baseband) was loaded with traffic. If the phase/amplitude characteristics were not very well aligned, intermodulation products from the voice signals fell in the baseband used by the digital signal and degraded performance.

All specifications on the T/S system were met in the design of the equipment. It is not known what margin was achieved, i.e., how much performance exceeded specification.

D. WU Initial Field Deployment, Testing, O&M

No unusual problems were encountered in field deployment other than those of the microphonic klystrons and the exacting phase/linearity requirements. The system was being operated as a test-bed, and engineers were used extensively to put it into operation. No legitimate measure of how well O&M craftsmen or other lower skill-level personnel could handle the system was obtained.

Digital test equipment was not widely available at the time, especially bit error rate test equipment, but these problems were not peculiar to the T/S system. Non-routine test equipment was not required for the T/S equipment, and in today's environment, one would not expect bit error rate equipment to present a problem.

In retrospect, no changes to the T/S system design which would enhance or otherwise favorably influence the field deployment and testing are known by the Western Union people.

E. WU System Operations Experience

The Western Union system did not go into regular commercial service so questions about operational experience generally are not answerable. Based on the several months of field trials, the system was judged to be performing as expected. Technical performance, reliability, and

maintainability are felt to have been as designed and no readily obvious change would be apparent. No flexibility for growth or change requirements have been identified. When queried about requirements now foreseen that would have influenced design, the Western Union engineers responded that the deployment and growth of AT&T's Digital Data System (DDS) would probably influence any such system they would design today. For interoperability reasons, and for cost/simplicity considerations, a design today would probably take a timing/synchronization reference signal from DDS at one or more locations and then distribute this over the Western Union network in the master-slave fashion. The Loran-C capability would not likely be implemented.

F. Other Information Useful to DoD Planning of DCS

The Western Union digital systems design/deployment experience did not result in specific T/S information other than that described above. An opinion was ventured by the Western Union engineers with regard to overall digital system, however: if possible, DoD would be well advised to avoid trying a hybrid system such as the digital MUX's on an analog radio. They feel that the most cost-effective, least troublesome route to follow today would be to initiate any such system as a pure digital network, and not attempt to incorporate existing hardware.

CANADIAN DATAROUTE

The Trans-Canada Telephone System (TCTS) is an association of eight of the largest telecommunications companies in Canada. Within TCTS, the Computer Communication Group (CCG) is dedicated to data communications services. After initial experimental work in 1971 with a synchronous digital network, CCG announced the development of a much more expanded synchronous digital network called the Dataroute. The system provides point-to-point service to fifteen metropolitan areas in Canada. It is presently in service carrying customer traffic. Its timing and synchronization system is basically a variation of the master-slave technique.

The Dataroute system was designed to utilize existing long haul facilities of the member common carriers of TCTS. Most of the long haul facilities of these carriers are 5 GHz microwave radio systems. Group band (12 equivalent 4 kHz voice channels) modems are the basic method employed in the Dataroute system to utilize these microwave facilities. These modems, used over the analog system, provide a 56 kb/s channel which is the basic trunk for the digital system. The modem typically interfaces with group level equipment in a FDM carrier system.

An alternate way of providing 56 kb/s trunks for Dataroute is by use of a Time Slot Access Unit on T1 facilities. This equipment converts the 56 kb/s signal to a 64 kb/s data stream suitable for insertion directly into a PCM voice channel. This type trunking is used in

Dataroute where T1 facilities are available.

Channelizing for Dataroute is provided by a two-level time division multiplex system. The highest level multiplexer is bit interleaved and synchronous on both the input and output. It provides channels for all circuits of 2400 b/s or greater speeds. It can be programmed for any combination of input channel speeds which are multiples of 200 b/s. The multiplexer in conjunction with the T/S subsystem minimizes delay through the network by time aligning all transmit and receive frames at a node. This capability to align the transmit and receive frames allows channels to contain submultiplexed channels which need not be demultiplexed as the composite channel is patched through a node on a drop-and-insert basis. This is achieved without adding framing overhead to multiplex the subchannels. This time alignment also allows patching data signals only (no clock) between multiplexers.

The second level multiplexer is normally programmed to accept both a bit clock and a frame clock from an external source (the T/S system). This configuration is used at nodes where a station clock is used - normally wherever two or more links come together. However, at terminal sites which connect to a single other location, the multiplexer can be programmed to slave to a bit clock and frame clock derived directly from the incoming line.

The lower or first level multiplexers used in Dataroute are character interleaved machines which provide asynchronous low speed channels. The high speed sides of these multiplexers are operated synchronously and feed into the low speed ports of the synchronous second level multiplexers described above.

From the beginning of the TCTS data communications network, the T/S system has evolved through three stages. Originally, a very simple master-slave system was used with a rubidium master located in Toronto. The next stage, and that which is primarily in operation now is a master-slave augmented with extra capabilities. This T/S system has been in operation since early 1974 in conjunction with the subsystems previously described and today forms the heart of the Dataroute. Dataroute people refer to the augmented master-slave system as Hierarchical Master Slave (HMS).

Plans are presently underway in the CCG to add a third level of multiplex to the Dataroute and with it to overlay a higher level T/S subsystem. This "new" T/S network will again be essentially a pure master-slave type which will obtain its reference from the same basic source as the present HMS system. A new third level multiplexer and the new T/S system will operate at 1.544 Mb/s and will provide channels for the existing 56 kb/s Dataroute. The two timing systems (new MS and old HMS) will organize and run independently except for the common

master reference.

The new higher level MUX and T/S equipments are presently in the early field trial and deployment stages and are not discussed further in this report.

The HMS system is discussed in what follows and is referred to simply as the Dataroute T/S system.

The Dataroute T/S system is implemented by a master-slave system in which double endedness and self-organization are incorporated. (Directed control is of course used as it is inherent in master-slave.)

The HMS system is like a conventional master-slave system in that network timing emanates from a stable master source located somewhere near the geographical center of the network (Toronto in this case). This timing is fed down the network in tree-like fashion. Self-organization is accomplished in a hierarchical fashion hence the HMS name. Accompanying the timing signal is a ranking, or figure of merit signature whose value at any location is generally dependent on how remote the location is from the master source. This signature is carried on a 400 b/s overhead channel. The signature is carried as a three digit number where digit one designates the node from which the clock first originates. Digit two signifies the number of links that have been traversed from the node having the original (master) clock, and the third digit carries the value of the immediately prior node. As the timing reaches each node and is passed on, the last two digits are updated. At every node in the network each incoming 56 kb/s stream carries such a signature. At a node with several incoming channels, the stream with the lowest valued signature is chosen as the one from which to derive timing reference. Continued monitoring will permit the network to automatically reconfigure in event of loss of master or a link failure; timing loops will not be set up in the reconfiguration. Automatic reconfiguration without setting up unstable timing loops was the primary reason for using the HMS system as opposed to a straight MS system.

Another system feature incorporated in Dataroute is called Master Frame (MF). This concept is basically the dissemination of a framing epoch marker by the timing supply at each node in addition to a bit epoch marker. In the 56 kb/s system, the basic frame is 280 bits long, and the frame epoch occurs at a 200 Hz rate. Thus the nodal timing supply, or station clock, delivers to each multiplexer both a bit clock at 56KHz and a frame clock at 200 Hz. This delivered frame clock is the Master Frame. This common frame marker is obviously time aligned at all multiplexers in a given station and is useful in minimizing delay through the network for a channel patched through a station. This minimization comes about because the low speed input buffer

requirements are minimized on each multiplexer. The Master Frame concept also allows drop and insert of submultiplexed channels without requiring overhead framing bits for the submultiplexing.

Another feature called Universal Time Frame is implemented in the Dataroute T/S system. Universal Time Frame (UTF) is an embodiment of double endedness. This embodiment actually involves an interfunctioning of the T/S subsystem with the second level multiplexers (56 kb/s). The designers of Dataroute included UTF to further minimize network user delays. They recognized that this feature would also allow precise time dissemination; however, it is not exploited in the system.

Universal Time Frame operates as follows. Consider the case where one of two communicating nodes is slaved to the other for timing dissemination purposes. The two multiplexers when initially put into service, will be programmed such that each transmitted frame is advanced in time (relative to local master frame) by an amount which slightly exceeds the propagation delay. This delay is known from theoretical or empirical (or both) considerations. Because the frames are advanced at the transmitter, they arrive at the far end of the link (nearly) in phase with master frame. The receiving multiplexer at the master station of the communicating pair measures the time alignment of the received frame by measuring the operating position of its high speed input buffer. This buffer position is then communicated to the slaved node via the overhead channel previously mentioned (with regard to the HMS signature information). At the slaved node, this information is used to generate an error signal to the local station clock. This signal causes the clock to adjust its frequency so as to advance or retard the time alignment of the frame marker. Correcting the slaved clock obviously adjusts its master frame epoch relative to the master station. This continuing, dynamic correction of the slaved station's master frame obviously serves to keep the slaved frame marker time aligned with the master station marker, and the alignment passes down through the network.

The transmittal of buffer position information to the other node constitutes a form of double endedness.

Specific answers to the set of questions follow:

A. <u>Fundamental Reasons for Choosing Dataroute T/S System as Implemented</u>

Efficiency (ratio of data bits to data-plus-overhead bits), ease of submultiplexing many different rates of user channels, and system simplicity were the chief reasons for choosing a synchronous digital network. The criteria for choice of which type synchronizing system were largely based on the Canadian network topology. It was envisioned that a long thin network would initially be deployed but that a highly interconnected topology would evolve. Flexibility to accommodate rela-

tively unpredictable growth was also considered essential.

The Dataroute T/S system is master-slave augmented by double endedness and self-organization. A straight master-slave system as well as a mutual synchronization system were also considered. The choice was made on grounds of enhanced security, ease of network expansion, and relative insensitivity to link failures. The choice was also tempered by a desire to retain simplicity to the extent possible. Thus performance was directly factored into the choice; costs and O/M influences were indirectly factored in.

Referencing the system directly to UTC or to a national standard was not considered as needed for Dataroute, but it was recognized that the Universal Time Frame feature allows distribution of (relatively) precise time. CCG did not foresee a need or market for this service to the extent that the additional hardware would have been cost justified. On the other hand, compatibility with other systems was felt to be a future need worthy of factoring into the choice of T/S system. It was expected that interfacing with other synchronous digital networks would be needed and it was recognized that a master-slave type system allows this simply by using a common reference as the primary master source.

B. Theoretical Design of Chosen Dataroute T/S System

After the choice was made as to the type of T/S system, a number of factors constrained the design of the hardware. This basic trunking medium was primarily to be over existing 4 GHz microwave systems and the design had to account for the characteristics of this transmission channel. The rate of delay variations due to the transmission media, the induced jitter on the timing signals, the probability of outages and link failures, and the probability/frequency of re-routing of transmission channels were among the factors considered.

Other factors which led to specifications on the T/S hardware included: the desired time to loss of bit count integrity in case of failures in the timing dissemination chain, the reliability budget for the overall system, and cost goals throughout the system.

C. Practical Design/Engineering Considerations for Dataroute

No significant obstacles were encountered in the design/development of the Dataroute T/S system hardware. Each node in the system contains two station clocks. Basically these clocks are made up of a highly stable VCXO configured in a phase-locked loop. The loop is designed with a very narrow bandwidth (nominally specified in microhertz) and with variable slew-rate control. Each clock also contains the logic circuitry to accept and compare up to 32 incoming reference signals plus signatures and the logic circuitry to compute its own outgoing signature. Additionally of course, there is digital frequency synthesizer circuitry for deriving the various clock and frame rate signals.

None of the above circuitry presented inordinarily difficult design problems. The logic speeds involved were not extreme and the usual precautions in board layout, wiring procedures, etc., were sufficient to ensure good performance.

D. Initial Field Deployment, Testing, O&M

Specific, identifiable installation problems were not encountered in the early field installation of Dataroute. As is described later, some difficulties were met that can generally be traced partially to the complexity of the method of implementing the double ended T/S system concept. These are discussed more fully below under Systems Operations Experience.

Because the deployment of Dataroute was planned to be carried out as a field trial, engineering-level personnel from both CCG and hardware contractor organizations were used extensively. The fact that the field craftsmen from the constitutent TCTS organizations were mostly experienced in FDM-FM analog technology influenced the decision to rely heavily on engineers for initial deployment.

No specialized or otherwise non-routine test equipments were required in the early field deployment of Dataroute. As with most digital systems, the indispensable tool utilized was an oscilloscope that could be externally synchronized. A logic analyzer type instrument designed to interface with the multiplexers, was used and is of course a specialized item. This however, was not necessarily required, i.e., the system could have been put into service without its use. More importantly, the use of this instrument was not influenced by the type of T/S system in use. The utility would have been the same with independent clocks, mutual synchronization, or whatever at the T/S method.

E. Systems Operation Experience

The Dataroute T/S system has operated as expected. Basic specifications have been met and no problems directly related to concepts or theory have been encountered. However, two issues have arisen. One of these relates to a practical problem on clock distribution in the terminal. The other relates more to system design philosophy.

The clock (and frame) signals in a Dataroute node are distributed from the station clock equipment rack to each rack of multiplexing equipment by means of a redundant bussing scheme. Two busses, one from each of the (redundant) station clocks, distribute signals to each rack of multiplex equipment. This distribution is accomplished via distribution modules mounted on the top of each bay (rack) of multiplex equipment. Several problems have arisen with this. It was found that although up to six multiplexers could be driven in tandem, this

was insufficient in some nodes and further expansion was not easily implemented. The lesson in this is to plan that the clock distribution system be easily expanded and that ideally, the expansion capability be unlimited in size.

Initially the clock signal was a current source-to-ground format. This contributed to the limitations on the number of multiplexers to be driven. It also gave rise to interference problems through noise pick-up, and common mode coupling. This problem was eliminated by redesigning with a format using a transformer-coupled, bipolar, balanced signal of several volts.

The second (philosophical) issue relates to the capability of the T/S system. As has been described, the Dataroute T/S system is basically a master-slave with three complementing features.

1. Hierarchial Master-Slave (self-organizing)
2. Master Frame clock along with bit clock
3. Universal Time Frame - a form of double endedness

In Dataroute, the reason for using Hierarchial Master Slave was to allow automatic reconfiguring of the network, without setting up unstable closed timing paths, in the event of link or master clock failures. However, experience has shown that network reconfigurations are so rare that the utility of HMS is questionable. In other words, the self-organizing feature is so seldom used that a cost/benefit analysis based on this experience would likely dictate not incorporating it in the design if the choice were to be made again for Dataroute.

Note: The rarity of failures has other implications for military systems. Because of the rarity, it is possible, even probable that personnel may be so "out-of-practice" that they cannot diagnose and correct T/S system problems in a timely manner and may actually compound the problem through incorrect activities. This of course could be disastrous in a wartime situation. Thus automatic self-correction may be much more desirable.

The second feature, Master Frame, was implemented to allow cross patching between multiplexers without the need for (low speed channel) input buffers. In retrospect, the Dataroute operators now question whether or not the ability to dispense with the low speed buffers is a significant advantage. With LSI and VLSI usage growing exponentially, the cost of adding buffering to channel cards is not large. The simplicity of channel cards is a positive feature however, and the ability to patch channels with only signal leads (i.e., without separate clock signals) is a positive feature. This reduces the problems that arise when large numbers of channels in a single rack make physical room for the signal and clock cables scarce and the opportunity for wiring err-

ors are compounded.

The third significant feature in the Dataroute T/S system is Universal Time Frame (UTF). The original reason for UTF was to minimize user delay through the network; UTF does this by assuring that the high speed frame alignment buffers in the synchronous multiplexers remain, on average and in absence of propagation delay variations, at their center or null positions. However, with the hindsight that has accrued through several years of operating experience, it is now felt by the Dataroute operators that user delay through the network is not critically important. This is increasingly true as data users begin to move away from protocols that employ acknowledgement type transmissions (such as Binary Synchronous protocol). Delay considerations are also less critical to users who continue to employ acknowledgement protocols but who take advantage of the better channels (lower error rates) increasingly available. Longer data block sizes are feasible (to a point) on the better channels and with the longer blocks/fewer errors, turnaround time, or delay, is less degrading to overall throughout.

The reliability of the T/S system on Dataroute has proven to be degraded by the complexity of the station clock. The HMS function plus UTF are implemented at the cost of increased complexity. This increased complexity serves to cause increased hardware failures plus contributes to maintainability problems as discussed below.

The maintainability of the T/S system on Dataroute has proven to be significantly more of a problem than anticipated. This is partially due to the complexity of the double ended, self-organizing system and partially due to the method of implementation. The system has been in operation over 5 years now and maintenance craftsmen do not yet totally comprehend the synchronizing hardware. Engineering level personnel are heavily involved in T/S system maintenance and operation. Even people with this level of training/competence lack in understanding the system. This complexity and the attendant lack of familiarity by the field people serves to reinforce the skeptical attitudes as to the cost benefits of UTF and HMS. Considerable effort has been expended to write and rewrite equipment manuals and to upgrade training, but as a practical matter, the maintenance people continue to regard the T/S equipment with considerable confusion. Much of these comprehension difficulties can be attributed to the specific method of implementing the double endedness. If, for example, the information passed from one end of the link to the other were available to maintenance/operation personnel in familiar units of time (seconds or microseconds), much of the difficulty would probably be overcome, (Note that the information is presented as a dimmensionless number referred to as "Diff" in the present implementation).

The flexibility of the Dataroute T/S system is judged to be accep-

table. No restrictions on growth or change to the network are imposed by the synchronizing system. It is notable however that in the planning for the 1.544 Mb/s network, which is a growth step for Dataroute, a straight master-slave technique for T/S is planned. Future requirements appear to be for a longer, less interconnected network - one wherein closed loops are less probable.

Advice to be offered to U.S. Government/DoD in planning for T/S of the future DCS was succinct: simplicity should be watchword. To the extent that survivability, security, and other specialized military considerations allow, the simpler the T/S system, the better will be the performance. Maintainability will be a strong function of simplicity and reliability/availability will probably re related to maintainability more than any other system parameter.

DATRAN DIGITAL SYSTEM

Data Transmission Company (Datran) was a specialized common carrier company formed in the late 60's to design, build, and operate a nationwide network for data communications.

The system designed and built by Datran is distinct from the others surveyed in this report because the total network was newly conceived, designed, and built as a digital system. There were no constraints of having to utilize existing plant or equipment; nor was the design forced into compromise in order to have a voice channel or other analog transmission capability. It was therefore feasible to consider techniques, including the T/S system, which perhaps were not feasible as candidates for the other systems surveyed.

The radio equipment in the Datran network is a digitally modulated, line-of-sight microwave system which uses 8-level PSK modulation. Fully loaded, it accepts two, phase-synchronous 21.504 Mb/s data signals as well as a coherent 21.504 MHz clock signal.

The timing signal is recovered and the data regenerated at every repeater in the Datran system. Timing recovery is accomplished in 14.7456 MHz phase-locked loops. The loop bandwidth (3 dB) is 1 KHz in the radio system. The free running stability of the oscillators in the timing recovery loops is $1:10^{-6}$ (short term).

A three level multiplexing system is used in the Datran network. All three Datran multiplexers are synchronous machines. Each utilizes bit interleaving to combine parallel low rate channels to a higher speed serial bit stream. Every multiplexer features buffers on both the high and low speed input ports for jitter reduction and to accomodate propagation delay variations.

The system which provides timing and synchronization for the Datran

network is a straightforward application of master-slave techniques. The T/S hardware is primarily made up of an equipment called the Datran Station Clock (DSC) which supplied bit rate clock to all levels of multiplexers. However, the overall network timing subsystem is composed of functions in the multiplexers, the DSC's, the microwave radios (as a timing dissemination channel), and one or more rubidium standards which provide the basic reference signal (master source).

The DSC basically is a highly stable, highly reliable disciplined oscillator with the ability to "remember" an input reference frequency after loss of the reference. It includes a frequency synthesizer for deriving various system clock rate signals from the basic oscillator frequency and line drivers for distributing these clock rate signals. It also includes logic hardware for interfacing with the multiplexers on a systems basis and for controlling its own functioning. The DSC provides clock signals at 21.504 MHz, 2.688 MHz, and 168 KHz.

The heart of the DSC is a triplicated set of modules called Timing Generators. These are highly stable 's, each in a phase-locked loop configuration where the loop error voltage is simultaneously fed to the VCXO and also digitized, filtered to an effective 0.04 Hz bandwidth, and stored in memory. The digitizing is done to a 12-bit accuracy which translates to a frequency accuracy at the VCXO of a 5×10^{-11}. This stored error voltage is applied to the VCXO in instances when the station clock transitions from slave operation to master. Because of the narrow bandwidth of the digitized error feedback loop, input jitter with frequency content down to much less than 1 Hz is filtered. By virtue of this jitter reduction, when the clock becomes a local master, frequency differences between it and the normal master will remain sufficiently small to maintain bit count integrity for more than 30 minutes (based on the buffers used in the C-MUX's). This is in the presence of up to 5 percent rms jitter with a bandwidth from 0.1 to 1000 Hz.

The bandwidth of the (undigitized) error voltage in the phase locked loop is 4 Hz. This translates to a capture, or pull in range for the Slave Node of $\pm 1 \times 10^{-7}$.

In the following, answers are given to the set of questions as they pertain to the Datran System.

A. <u>Fundamental Reasons for Choosing Datran T/S System as Implemented</u>

The criteria used for choosing the T/S method for the network were principally performance, cost, and state of the art (practical state as opposed to theoretical) prevailing at the time of system design. Methods other than master-slave were considered. Nonsynchronous operation using independent clocks and various types of pulse stuffing were evaluated as possible candidates. Synchronous techniques considered

included designated (or independent) master, mutual synchronization, and of course master-slave. Bit stuffing was ruled out as being too expensive to implement in a dedicated circuit system. In a switched system, the cost is even more prohibitive. Independent clocks were discarded as a potential candidate mainly because the prospect of periodically overflowing buffers, no matter how great the period, was judged to be a severe marketing disadvantage. Of the types of synchronous operation considered, master-slave was chosen for cost, performance, flexibility and simplicity reasons. The Datran network was envisioned, even far into the future, to be a topologically long thin system with no looped subnetworks. Because the basic transmission medium at 44 Mb/s had extreme capacity in terms of customer channels, it was not considered likely that alternate routes and possible loops would be necessary for many years. Too, the transmission medium was to be designed for such reliability that the need for failure-induced alternate routing and the possibility for looped configurations was diminished. Last, since transmission was to be entirely over Datran facilities, routing configurations could be tightly controlled. All these reasons enhanced the master-slave choice, without self-organizing complications.

Simplicity of hardware and the attendant advantages in maintainability and reliability also weighed in favor of master-slave as compared to frequency averaging (mutual). Independent masters appeared to be a viable choice as regards simplicity and cost (assuing an existing system such as Loran-C as the basic master source). However, the degree of dependency on a system whose operational reliability and integrity lay outside the Datran operation was considered a negative aspect. And significantly, in the case of Loran-C, it was felt that the availability of ground wave coverage was lacking for the total geographical area Datran hoped to eventually serve.

Simulation was not used in choosing the Datran T/S system. Formal performance and cost analyses did influence the decision; however, engineering judgments on practicality and probability of successful, on-schedule design also weighed heavily.

System reference to UTC or other standards was not considered a strong factor in evaluating T/S techniques. It was recognized that master-slave afforded relatively easy interfacing with one other synchronous system (by utilizing either a common reference source or by having the Datran network be a simple extension of the other network), but again this was not a strong factor in the T/S choice.

B. Theoretical Design of Chosen T/S System

The Datran system was designed "starting from scratch" and choices were not constrained by existing plant or facilities. LOS microwave transmission media was an early choice. The yearly (or longer) and

daily delay variations over multihop, multiclimatic microwave chains was not precisely known, and the T/S system chosen would have to be versatile enough to handle this unknown.

The following specifications were design goals for the Datran T/S system:

a. Stability (without primary reference)

   $\pm$ 1 x 10^{-9} per day

   $\pm$ 4.5 x 10^{-8} per 6 months

   With a jitterless primary reference input, the T/S station clock was specified to have less than 1 percent rms output jitter.

b. Reliability

   The station clocks in the Datran network were specified to have a mean time between catastrophic failure of greater than 1,613,000 hours.

c. Jitter Reduction

   The station clocks were specified to operate satisfactorily over up to 33 tandem microwave links. To accomplish this, a jitter reduction of at least 6 to 1 (for input jitter contained in a 0.1 to 1000 Hz bandwidth) was specified.

d. Cost Goals

   In 1974 dollars, cost goals in the range of $50K per (fully configured) station clock were established.

e. O/M Aspects

   Guidelines on modularity, number and nature of visible alarms and indicators, and diagnostic aids were established. A mean-time-to-repair (after reaching equipment site) of less than 30 minutes was established as a design goal.

The principal criterion was performance as opposed to conformity with existing standards. This was especially true for interfaces between multiplexers and between multiplexers and the radio. At customer interfaces, conformity with widely used standards was necessary. At all interfaces in the Datran network, the timing signal is carried on a separate conductor from the data signal. This contrasts with the possibility of having the clock signal not be distinct from the data signal as is done for instance in the familiar bipolar of T1 systems.

C. Practical Design/Engineering Considerations in Datran T/S

Few notably difficult or unexpected obstacles were encountered in the design and development of the Datran T/S hardware. The design of

majority voting logic for controlling the triplicated functions in the DSC required close attention to signal delays, board layouts, etc. Adherence to good engineering practices common to nanoseconds speed logic design ensured a minimum of problems in this area.

One problem was encountered in a systems level test of the T/S equipments. The problem related to the ability of the DSC to properly hold the system frequency when a jitter corrupted reference signal was lost. It was found that the frequency spectrum of the jitter on the reference was of critical importance in how well the system frequency was maintained. The overall specification required that when using 32 bit buffers, bit count integrity would be assured for at least 30 minutes when a reference with 10 percent rms jitter, with a frequency spectrum between 0.1 and 1000 Hz, was lost. Recall that the bandwidth of the loop in the DSC which held the error voltage, or "remembered" the system frequency, is 0.04 Hz. However, it was found that if the jitter spectrum was concentrated near the lower end of the specified bandwidth, the loss of reference would sometimes leave the remembered frequency deflected enough to cause buffer overflow in less than 30 minutes. If the jitter spectrum were spread uniformly over the 0.1 to 1000 Hz band, or concentrated near the high end however, the specification was easily met.

This problem of maintaining bit count integrity in the presence of low frequency jitter was never manifested in actual field operation but only in lab tests where jittered references were synthesized.

With the exception of the above described phenomenon, all specifications on the T/S system were met or exceeded by the completed design. For example, a jitter reduction ratio of greater than 20 to 1 was measured in the field for the DSC whereas the specification was 6 to 1. A mean time between maintenance repairs of greater than 22,000 hours was also measured in the field on the T/S equipment. This compares to a design calculation of 6700 hours. As for mean time to catastrophic failure, no such failures had been encountered at last check so that a meaningful quantitative statement cannot be made.

D. Datran Initial Field Deployment, Testing, O&M

In the initial deployment and operation of the Datran network, all field personnel were given comprehensive, but extremely time-compressed training on all equipments. By and large, the personnel were experienced in FDM-FM analog systems as opposed to digital systems; thus many of them encountered and progressed through a distinct learning period on the entire network. While the T/S subsystem presented no more problems during this learning period than did the other equipments, nevertheless the nonfamiliarity was of significance. The triple redundancy of the Station Clock added to the nonfamiliarity problems to such an extent that it considerably complicated the understanding of the equip-

ment. It also greatly added to the difficulty of troubleshooting equipment malfunctions.

Datran procured several items of test equipment related to the T/S system that were nonroutine. Among these were portable rubidium standards, frequency synthesizers, and phase jitter meters. In retrospect, it is apparent that these equipments were not necessary for operating and maintaining the network. They were useful in monitoring the network performance from the point-of-view of an extended field trial where continued engineering data was desirable. However, from a standpoint of routine, day-in, day-out system operation and preventive maintenance, an oscilloscope in conjunction with built-in test functions on the DSC is adequate.

No significant changes in system design are obvious in retrospect that would enhance or otherwise favorably influence field deployment and testing.

E. Datran Systems Operating Experience

The T/S subsystem in the Datran network has operated substantially as expected. No significant problems have been encountered that would give reason to reevaluate the basic concepts or upgrade the system. However, it is probable that a simpler, less complex system would be designed if the task were to be done again. Some reasons for this are noted below.

The fully triplicated, highly reliable (and relatively expensive) DSC was deployed by Datran at every three-way junction station and at every terminal. In retrospect, such reliability and expense is probably only justified at remote three way stations. Even there, the need for triplication is questionable. The DSC's provide quite long mean-time between maintenance repairs; and the nature of failures encountered to date indicate that double redundancy (and the simplified control hardware inherent in this) would have sufficed.

Likewise, experience gained in actual operations of the system tended to indicate that early concerns over jitter on the recovered timing signals was not warranted. In the actually deployed Datran network, the longest span between DSC's was eleven microwave hops. Had the system been deployed to the west coast as planned, spans of over thirty hops would have been encountered. Nevertheless, experience with the eleven hop system indicated that jitter buildup was much less than expected. The worst measurement for the eleven hop system was jitter of less than 4 percent rms which would extrapolate to much less than 10 percent jitter on a 30 (+) hop span.

The Datran network actually deployed was a topologically thin, long line network with no closed circuit configurations. Thus no opportun-

ity ever arose to evaluate bit count integrity type questions on loss of continuity of the timing dissemination chain.

Performance of the Datran T/S system was judged adequate, and reliability was more than acceptable and perhaps over designed in some applications.

The maintainability of the T/S system probably should be judged acceptable to slightly marginal. The only questionable area would be in the use of triplicated functions as noted above. This tended to obscure craftsman understanding of the equipment and to render trouble shooting more difficult. It was certainly not a severe problem particularly for the better qualified, more industrious field people who would put forth the effort to understand the hardware.

The flexibility for growth/change of the Datran Station Clock is judged acceptable. It was found that in one instance, added driver ports were needed to enable clocking additional multiplex bays in a station. Careful, tedious work was necessary to add the capability to an on-line DSC without disturbing customer traffic, but it was accomplished. Should a redesign ever occur, it would be desirable to provide for easier (physical) expansion of the drive capability; this was not a major problem, however.

F. Any Other Information Useful to U.S. Government DoD in Planning Defense Communications System

The same advice regarding simplicity of equipments as offered by Canadian Dataroute operators would be evidenced by the Datran experience. The less complex the system, the better will be the reliability and maintainability. And, in the absence of major design errors, maintainability will probably impact performance more than any other single factor.

One other comment is pertinent. Compared to the late sixties/early seventies when the Datran network was designed, there is a plethora of data on digital systems available now. Each succeeding conference or symposium on broadband communications provides additional experience documentation (as well as theoretical work). Notable useful examples are LOS microwave delay data and multipath fading degradation experience. Obviously, all these sources should be used in planning the future DSC timing and synchronization systems.

TIMING/SYNCHRONIZATION PLANNING FOR NO. 4ESS

As noted in the introduction, original plans were to include information on the implementation and operational experiences of Bell System's Digital Data System in this report. (This system is a private line synchronous network put into operation in the U.S. in 1974). Unfortunately, this information is not available; Bell officials decline

to provide such data due to the extensive pending litigation revolving around the Digital Data System (DDS). These officials were, however, willing to provide information on the current planning and engineering efforts related to synchronization of a network of large digital switches presently being implemented throughout the U.S. This information was provided by way of interview with several engineers currently involved in the engineering and planning efforts. In this section, this information is described.

The No. 4ESS is an electronic, software-controlled toll switching center designed to tie local facilities to the nationwide long distance network. The switching center is basically a digital machine which interfaces both digital and analog trunks. At the inception of No. 4ESS deployment, and at the present time, analog trunks predominate in number. As digital trunks become increasingly available however, (a trend that is growing rapidly), the interconnected No. 4ESS systems take on more and more attributes of large integrated digital networks. As this happens, the advantages of reliably synchronizing the whole network are obvious. This is recognized by Bell and considerable effort has been expended to plan for this synchronization. Much of this planning is inextricably interwoven with planning and analysis that has been going on at Bell Labs since the early sixties. This early planning was the initiation point in the discussions with Bell engineers.

In the early sixties, various users of the Bell Telephone System began to implement digital networks. At first, relatively elaborate word stuffing was used to provide a synchronizing mechanism for the networks. The cost (at that time) of shift registers caused a change to bit-by-bit pulse stuffing techniques and eventually to consideration and analysis of other synchronizing methods. Considerable attention was focused on the concept of mutual synchronization. Many investigators, both inside Bell Labs and elsewhere, studied and published papers on the technique. One of the advantages ascribed to mutual synchronization was that it was "administration free". However, it was decided by the No. 4ESS network planners that this is not, in reality, the case. As an example, it was noted that up to 20,000 DS-1 (1:544 Mbps) trunks can converge on a No. 4ESS. To choose which of these would provide the timing reference paths for a mutually synchronized system is a considerable problem. This is especially true when the dynamic nature of the topology of these trunks is considered.

Another aspect of mutual synchronization noted as significant was the dumbell effect. This situation arises when two separate, relatively complex, interconnected areas are connected by a thin trunk structure. In this case, the input to the T/S systems in each of the ends of the dumbell must carry heavily weighted importance. This tends to distort the survivability of the mutual sync network and to compound administration problems, particularly when the exact composition of the

"thin" trunking is dynamic.

Mutual synchronization was eventually discarded as a candidate synchronizing method for three significant reasons. As indicated, it was determined that its main attribute, i.e., administration-free operation, was not a practical reality. This was the primary reason for rejecting the technique. Also, evolution in crystal oscillator technology in the late 60's was a factor in moving away from mutual synchronization. Double oven crystal oscillators with millidegree temperature stabilities became economic realities. The resulting frequency stabilities added to the case for master-slave techniques. Additionally, digital phase-locked loop (DPLL) technology progressed to the point that economic practicality was achieved. The significance in this was that DPLL's with "automatic" integral-plus-proportional feedback was then achievable. Note that integral plus proportional feedback had long been used, but the integral feedback had been a mechanical screwdriver adjustment (to oscillator free-running frequency). "Feedback" of course was a technician periodically nulling loop stress voltages. Digital implementation of this function meant greater automatic long-term stability in the oscillators since drift due to crystal aging could be compensated.

Another influencing factor on synchronization planning was that the DDS network began to be engineered. At first, the objective was to build a totally slip free network for DDS. However, the No. 4ESS planners took the position that a totally slip free DDS was not practically achievable. As an example of the several reasons, most of which are of an administration/operation origin, it was pointed out that there are some 60,000 separate group channel trunks in the Bell System. Of these, an average of 20,000 yearly are dynamic in the sense that they are removed and subsequently returned to service for one reason or another. Since these trunk assets were to provide the transmission channels for DDS, it was felt that such a dynamic makeup would prohibit practical operation of a slip free network.

For the DDS, it was eventually agreed that three system impairments would be foremost. These were i) bit errors; ii) misframing where multiplexers lose frame synchronization and iii) slips where single eight-bit bytes are lost/gained. Additionally of course, timing jitter due to accumulation on the transmission lines and due to pulse stuffing/destuffing is encountered. Of these impairments, the most significant in terms of timing and synchronization is slip and the byte interval defining slip. Of course, the longer the byte interval, the less the frequency accuracy requirements on a timing source for given slip rate. An upper limit is placed on byte length in the Bell System because the byte buffers give rise to cross-office delay in patching channels. At a point, this excessive delay begins to cause echo suppressors in the circuits to hang. This led to the 8-bit byte which is 125 microseconds in a 64 kb/s channel.

After the byte length was determined, an acceptable slip rate was evaluated. A number of factors were considered by Bell in determining a slip rate specification. In voice channels slips are trivial. However, for voice band data channels, slips can result in modems retraining which results in outages of up to several seconds. Slips also can have a severe impact when Common Channel Interoffice Signalling is carried over a slipping channel. The effect of slips of digital data varies depending on the application. It was noted by Bell that commercial applications of secure voice are foreseen over data channels in the future and slips could result in cryptographic sync problems.

The above and other considerations led to the Bell System eventually settling on a specification of one slip per 5 hours for an end-to-end reference circuit. This translates to a relative clock frequency difference specification of 1.7×10^{-9}. Because the possibility of longer frames becoming desirable in the future however, a tighter stability specification of 1×10^{-10}/day was agreed upon.

At the present time, the working specifications of the timing and synchronization subsystems for the No. 4ESS network are:

1. Clock frequency inaccuracies of less than 1×10^{-10}/day.

2. Synchronizing unit at each switch will phase lock to an external reference. This reference will be either a DS-1 signal at 1.544 MHz or the standard Bell System Reference Frequency (BSRF) 2.048 MHz. Note that this latter reference is presently distributed throughout the network.

3. The synchronizing system will be arranged in a hierarchical structure with frequency (not timing) information emanating from an atomic standard at Hillsboro, Missouri.

4. The timing system will have a highly structured maintenance plan administered out of a centralized location.

Note that specifications 2 and 3 define a master-slave system when the DS-1 reference is disseminated throughout the network. When the BSRF is used, an independent master system is being implemented.

It was noted by the Bell engineers that the fourth entry in the above specification list is extremely important as contrasted to its innocuous appearance. It has historically been found that maintenance activities are always reflected negatively in reliability performance figures for communications networks. Significant percentages of outages are traceable to maintenance personnel activities. Since the timing sources for digital networks are so central in the operation, and outages are potentially so catastrophic, it is felt that the structured

maintenance philosophy is critical. In general, alarming and alarm reaction plans for the timing and synchronization hardware on No. 4ESS will have long time constants. A loss of reference alarm will, for example, be considered only a minor alarm for up to several minutes. Other reaction times numbered in hours are planned. The central administrator with network-wide status visibility will be consulted closely on all alarm responses and maintenance activities.

The phase-locked loop in the TSU incorporates integral-plus-proportional feedback. The integrated component of the loop feedback has a time constant of approximately 2 days. The time constant of the proportional factor is approximately 2 hours. Additionally, a manually activated Fast Start mode allows rapid acquisition when the TSU is first put into operation or following a major reconfiguration or outage. Automatic phase buildout is also designed into the TSU hardware.

The above briefly describes many of the factors which have entered in the planning for No. 4ESS timing and synchronization. The Bell engineers emphasized that many of the choices and trade-offs made were dominated by network administration considerations. This of course reflects the culmination of many years experience in operating the largest common user communications network in the world. Little or no consideration has been given to some factors important to military strategic communications systems. Survivability in the presence of hostile activities is an important consideration in evaluating timing and synchronization techniques for the DCS. Understandably, this has had little or no influence on the Bell designs. Likewise, interoperability with tactical systems is a consideration for the DCS not dealt with by Bell.

A last topic discussed with the Bell engineers during the interview concerned the distribution of precise time. The T/S system being engineered for the No. 4ESS network does not incorporate double endedness and thus does not readily admit of precise time dissemination. However, consideration has been given to this by Bell engineers. In general, the same concern for network administration considerations influences time dissemination that influenced T/S system design. In fact, the Bell engineers expressed the opinion that administration problems would so adversely effect time distribution as to make it impractical in the Bell System. Two examples noted were network topology dynamics and emergency restoration. The statistic previously described wherein 20,000 of the 60,000 group trunks in the system are taken out of service and restored each year was pointed to as an example of the dynamics.

Emergency restoration, and the tempo of activities that exist during outages was mentioned as the other prominent example of administrative problem in time dissemination. During outages, especially those that

effect large cross-section trunks, intense effort is involved. Customers and management are usually demanding quick restoral and little acceptance of procedures which would inhibit or delay restoration would be anticipated. Coordination of clocks, and of new trunk path delays, etc. is seen by Bell engineers as just such a procedure. It was noted however, that automation of such procedures by use of microprocessors might become economically feasible in the future.

SUMMARY OF SURVEY

Each of the systems surveyd is a synchronous network, and each uses a T/S technique that is essentially of the basic master-slave category. The Datran network T/S system and the planned No. 4ESS netowrk T/S are straightforward applications of master-slave. The Western Union network and particularly the Canadian Dataroute network T/S are master-slave with added features.

Retention of bit-count-integrity, i.e., the avoidance of slips was the foremost criteria in the design of the networks. Flexibility and adjudged practical state of the art in T/S technology also weighed heavily in the initial system designs - particularly in the case of the Western Union and Datran networks. Minimization of user delays over the network was a strong factor in the Dataroute system design; recognition of network administration problems strongly influenced Bell System plans and choices.

Only the Canadian Dataroute system provides for automatic network reconfiguration (and automatic network initialization) following failure or other perturbations. Related to this is the fact that among the four systems surveyed, only the Dataroute requires an overhead channel.

Each of the system designers recognized the crucial importance of realibility in the T/S hardware/subsystem. Outages due to lack of timing supply are potentially catastrophic and could disrupt very large numbers of users. Thus, considerable effort and dollars were expended to build high availability into the T/S equipment. Western Union and Canadian Dataroute system architects accomplished this by using double redundancy with automatic switchover. The Datran system utilized triple redundancy in the station clock with all three portions contributing in the unfailed, or normal, operating mode. Majority voting logic in effect switches out the failed circuitry when problems arise.

The field deployment experiences of the organizations surveyed were remarkably similar. Each of the Western Union, Dataroute, and Datran initial deployment efforts were reasonably straightforward, but extensive use was made of engineering-level personnel (as opposed to technicians or craftsmen). The lack of a ready supply of digitally oriented field personnel was noted. In each case, however, the T/S system

type or its features had little or no influence on deployment difficulties or lack thereof.

In the area of operations experience, only Datran and Dataroute were able to contribute information. The Western Union network never went operational and the Bell System's deployment of No. 4ESS is only now underway. The Dataroute operational experience has been adversely effected to some extent by the conceptual (and hardware) complexity of the T/S system and to implementation procedures. To date, field personnel are not totally comfortable with their knowledge and understanding of the method of timing and synchronizing the system with the extra features of the master-slave system. Notably, the next evolution of the Dataroute system is planned to be an unadorned master slave.

In the case of Datran operational experience, no significant problems arose which were attributable to the T/S concept. However, the triple redundancy in the hardware did complicate craftsman understanding and troubleshooting. Failure rates were so low as to call into question the need for triple as opposed to double redundancy.

The above briefly summarizes the information gained in the survey of commercial digital networks. It should be noted that several key criteria which greatly influence the design and (satisfactory) operation of a strategic military network such as the Defense Communication System are absent from the list of commercial network criteria examined. Among these are:

1. Survivability
2. Compliance with Federal Standard 1002
3. Intersystem Operability
4. Self-reorganization
5. Precise Time Availability

In the above list, survivability is the most significant entry; the fact that a network plays a central role in national defense means that survivability must be as pivotal in design trade-offs as are any other parameters. This, of course, can completely alter the conclusions of a cost-effectiveness study.

Precise time availability is another feature that might be of considerable value in a defense communications scenario but of lesser importance in the commercial world. Covert communications requirements lead to spread spectrum and cryptographic system use. Precise knowledge of a coordinated time standard, or course, greatly facilitates the synchronization of such systems by reducing acquisition search windows. Likewise, a knowledge of relatively precise time facilitates navigation and position fixing - another distinct advantage in a military environment but of lesser importance for commercial applications. (Although the commercial need for precision navigation

aids should not be minimized.)

The commercial networks surveyed have made no attempt to use their T/S systems in any manner to augment or form a basis for network monitoring and evaluation. This is of course an area where the benefits could be of importance to a commercial system even though it doesn't have the special requirements of a military system. Such a use of the T/S system would generally require preplanning and engineering of the monitoring and evaluation capability into the hardware during initial design, something obviously not included in the systems surveyed. Recognition of the benefits to be obtained through monitoring and evaluation could influence the type of T/S system chosen. Some T/S methods are more compatible and synergistic with monitoring and evaluation than others.

A concluding comment is in order. Simplicity and administration related issues were mentioned over and over by the systems operators. It was emphasized that these aspects should be central in the design of any T/S system. The reasoning behind these statements is not presently debatable, especially among experienced systems operators. However, the potential influence that microprocessor evolution is having and will continue to have on this aspect should not be overlooked. Increasingly, it is feasible to automate activities heretofore constrained to be manually accomplished. With automation, or machine driven activity, the rationale behind simplicity arguments should be reexamined. Without the factor of human error, much in the way of automatic corrective routines are feasible. And much in the way of system features now considered very useful, but too complex to be practical, become realistic. This can and certainly should be factored into any current T/S systems design.

QUESTIONS AND ANSWERS

MR. THEODORE FRENSCH, Hewlett-Packard:

Can you elaborate on the double or triple redundancy? I can see for catastrophic failures that double redundancy works where, simply, there is no output. But what is the practicability for frequency shifts in double redundancy?

MR. MITCHELL:

That consideration is one of the things that led to triple redundancy choice in DATRAN, the idea being that for amplitude failures, certainly double redundancy allows you to decide which one has failed and which one is operating normally so you can switch appropriately. But for phase-type changes, if you have a phase shift in one of two sources, it is difficult if not impossible to determine which one is right and which one is wrong. Whereas, if you have three, you can do majority voting.

That is well and good. Our experience, again from a practical point of view, was the clock as it was built. Admittedly, we spent an appreciable sum to make it reliable. But it was so reliable that that came into play so infrequently, never in my experience with DATRAN, that that almost fell into the realm of an academic consideration. Does that answer your question or address what you are driving at?

MR. FRENSCH:

Are you saying, then, you actually don't need any redundancy or are you saying that catastrophic failures were more common than frequency shifts?

MR. MITCHELL:

Yes, in our case we didn't have either. When you have a network that potentially has 8,000 channels on it, you can't afford to have no redundancy, even if it only happens once every 30 years. First of all, if you had to get out and try to market the network and somebody started asking questions about the market service on the network, it would cause you problems.

In fact, some of our station clocks in the three-way repeaters were 60 miles from the maintenance center. So you are talking about hours before the service could be restored if you didn't have some kind of backup there and the clock did fail. So I think you can't afford to go with no redundancy just from a practical point of view.

But as far as going to the expense, and I would guess that the expense of the DATRAN clocks went up 50% due to the triple redundancy, and certainly the development problems were increased by at

least 50%; you know, the design problems for the people that provided it. Had they not had the triple redundancy to contend with, the design would have gone a lot more smoothly; not to say that it didn't go relatively smoothly anyway. Yes, sir?

DR. C. C. COSTAIN, National Research Council:

Data Route has installed a terminal in our laboratory and we continuously monitor their frequencies. We are, in fact, going to feed back to them digitally on the network their offset, which they can use as they wish operationally. But I just wanted to mention that after we were monitoring, there was a catastrophic failure of their master in Toronto, or their sync channel, whatever, or something in the loop. Montreal took over and they were rather pleased they had it because it was a holiday weekend with, of course, the least experienced, most junior engineer in charge. The system did exactly what it was supposed to do in case of a failure of equipment. We monitored it beautifully and we were rather pleased to have the experiment, which you wouldn't create normally.

MR. MITCHELL:

Very good. You touched upon a point I neglected to mention that did come out in the survey. The fact that these timing and synchronization systems seldom fail can be important, particularly in a military-type environment, because it means that the people who have the responsibility for maintaining them when they do fail, have forgotten what to do because it's been so long since they had a similar problem. So if you have a network that automatically heals itself in some sense, that is an advantage. It is probably even more so an advantage in a military network or critical network than it would be in a commercial network.

Page intentionally left blank

AN EVALUATION OF OPTIONAL TIMING/SYNCHRONIZATION FEATURES TO SUPPORT SELECTION OF AN OPTIMUM DESIGN FOR THE DCS DIGITAL COMMUNICATION NETWORK

D. B. Bradley, J. B. Cain, III, and M. W. Williard
Harris Corporation, Melbourne, Florida

ABSTRACT

This paper reports the results of one of the tasks of a study performed by Harris Corporation for the Defense Communications Agency (DCA). The task was to evaluate the ability of a set of timing/synchronization subsystem features to provide a set of desirable characteristics for the evolving Defense Communications System (DCS) digital communications network. The set of features relate to the approaches by which timing/synchronization information could be disseminated throughout the network and the manner in which this information could be utilized to provide a synchronized network. These features, which could be utilized in a large number of different combinations, include mutual control, directed control, double ended reference links, independence of clock error measurement and correction, phase reference combining, and self organizing. Some additional secondary features include smoothing for link and nodal dropouts, unequal reference weightings, and a master in the mutual control network. The set of desirable characteristics used in the evaluation of the features includes but is not limited to, frequency and phase accuracy, minimized propagation of disturbances in the network, slip free communications as a normal mode of operation, survivability under stressed conditions, interoperability with other networks, monitorability, precise time availability, minimized overhead requirements and cost effectiveness.

The utility of each feature (and combinations thereof) in providing the desirable characteristics was evaluated by means of combined analysis and computer simulation. One underlying assumption was that of a microprocessor based nodal synchronizer which implemented a second order digital phase locked loop with an extended range linear phase detector. The nodal microprocessor was also utilized in distributed control of the timing/synchronization subsystem network.

Some of the features of the simulation model included: 1) a single network topology of 20 nodes and 29

interconnecting links (a mixture of cable, satellite, microwave, and tropo), 2) normal variations which included an initial startup transient, normal path delay variations, natural clock frequency drifts, and link and nodal measurement jitter, 3) three sets of stress scenario and 4) sixteen combinations of the timing/synchronization subsystem features.

Nodal loop parameters used in the simulation were not optimized but were a compromise of several conflicting goals. These included maximizing utilization of the available reference information, providing for rapid entry into the network, coasting through reference outages and minimizing the effects of transient and steady state reference perturbations. A continuing portion of the study from which this paper was abstracted aims at optimizing these parameters.

The relative abilities of each of the sixteen combinations of the subsystem features in providing each of the desirable characteristics is presented in tabular form for easy comparison. Although this study is particularly concerned with the DCS application, the information developed could be applied to any extensive network.

I. INTRODUCTION

A large switched digital communications network gives rise to a set of desirable characteristics for its timing/synchronization subsystem. This set of desirable characteristics is identifiable from the mission of the communications network. There are a number of features which may be applied to the timing/synchronization subsystem in various combinations to provide the identified set of desirable characteristics to varying degrees.

This paper reports the results of one of the tasks of a study performed by the Harris Corporation for the Defense Communications Agency (DCA). The purpose of the task was to evaluate the ability of several timing/synchronization subsystem features to provide a set of identified desirable characteristics for this subsystem. Although the study was performed specifically for the future DCS digital communications network, many of the findings are also applicable to civilian digital communications networks. The approach used in the evaluation was combined analysis and computer simulation.

The set of identified desirable characteristics for the future DCS digital communications network timing/synchroni-

zation subsystem are discussed in some detail in [1]. For this reason, and since they are somewhat particular to the DCS network but primarily self explanatory, the set of desirable characteristics are merely listed in Section II. Since the set of features range from the exceedingly familiar to the possibly never-heard-of, a fairly detailed explanation of what they are is given in Section III. Attributes of the simulation model are given in Section IV. The results of the evaluation are discussed in Section V.

II. THE DESIRABLE CHARACTERISTICS

The set of desirable characteristics against which the features were compared included the following:

1. Frequency and phase accuracy.
2. Freedom of clocks from being disturbed by perturbations occurring further from the master than the local node.
3. Freedom from propagation of clock errors except that of the master.
4. No harmful propagation of a disturbance due to path delay variations or link dropouts.
5. Slip free as a normal mode of operation.
6. Survivability of the timing function.
7. Minimum overhead communications.
8. Precise time availability.
9. System level monitorability (functional as opposed to equipment monitoring).
10. Compliance with Federal Standard 1002.
11. Cost Effectiveness.
12. Interoperability of the digital communications system with other digital communications systems employing different synchronizing techniques.
13. Capability to automatically select a new network master.

III. THE FEATURES

There are several optional features which can be included in various combinations in the DCS timing/synchronization subsystem. Some of the features, however, preclude incorporation of certain other features. Also, some combinations of the features are more synergetic than others.

1. Independent Clocks (IC)

Figure 1 shows the form of this feature. This feature precludes the attainment of a truly synchronized com-

munications network due to the unavailability of ideal clocks. However, it provides a basis for many comparisons and it is also a natural fall back configuration for the directed or mutual control features, when in times of stress, all references are temporarily lost. It is spoof proof, its set up time minimal, and for a tactical military communications mission it may very well be the best choice. It is conceptually simple from the standpoint of the timing subsystem itself but it is not necessarily the simplest from the standpoint of the digital communications equipment designer or user, because lack of a truly synchronized network implies occasional resetting of communications buffers, as well as an increasingly complex system of pulse stuffing to accommodate growing multiplexing and switching needs. For a given quality of communication it is not necessarily the lowest cost system due to the higher implied quality of nodal clocks. Also, to the degree that survivability is reflected in the probability of bit slips and the ability to monitor the system for impending failures, it does not necessarily provide as survivable a network as certain combinations of the other features, all of which are precluded by independent clocks. Since independent clocks do not lead to a synchronous network it was not included in any of the simulations and will not be further considered in this paper.

2. Directed Control (DC)

The form of the directed control feature is shown in Figure 2. It consists of a tree network of selected communications links over which timing information is passed from master node to all immediate neighbor nodes. These immediate neighbor nodes in turn pass timing information from their nodal clocks to their immediate neighbors who are farther removed from the master. The timing information only flows in one direction of the duplex links, i.e., away from the master. This is the central idea of a master slave network which is a particular case of directed control. Since there are no closed loops in the network, this feature has no network stability problems. It is a widely used feature in military and civilian networks [2,3,4]. It results in a synchronized network in which all nodes have the same average frequency. When combined with elastic communications buffers of sufficient size, short term perturbations will not cause bit slips. Thus the network should be able to support communications without slips or scheduled resetting of communications buffers for an indefinite period of time.

3. Mutual Control (MC)

Figure 3 shows the form of the mutual control feature. It is the antithesis of directed control. Each node takes a weighted sum of the phase error of its local clock relative to that of signals received from all of its immediate neighbors to determine a correction signal to apply to its local clock. Since there are numerous closed loops in the network, care must be taken in the selection of error signal processing parameters at each node in order to insure network stability. The network frequency under mutual control is a function of the weighted average of the individual nodes natural frequencies and the path delays of the network. This feature has been widely studied from a theoretical standpoint, mainly to investigate the questions of network stability and the sizes of transient and steady state phase errors in the system [5,6,7,8,9,10,11,12,13,14] but no major network has selected this feature. It's most touted attribute for a military mission is its survivability.

4. Double Ended Reference Links (DE)

Figure 4 shows a model for double ended reference links. This feature provides the capability to remove the effects of transmission delay time in the time reference information. Each node in the timing hierarchy transmits to each of its immediate neighbor nodes its clock reading as well as the difference between its clock and that received from its immediate neighbor. This is sufficient information to remove the propagation time from the reference information. To see how this works, let $K_A = T_A - T_B + D_{BA}$ be the difference between node A's clock reading and that of node B. Notice that the propagation time from B to A has been explicitly accounted for by the term D_{BA}. Similarly, let $K_B = T_B - T_A + D_{AB}$ be the difference between the reading of node B's clock and that of node A with propagation time from A to B explicitly expressed by D_{AB}. Now node B may determine the error of its clock relative to that of node A to within one-half the asymmetry of the duplex link propagation times by computing

$$\frac{K_A - K_B}{2} = (T_A - T_B + D_{BA}) - (T_B - T_A + D_{AB})$$

$$= (T_A - T_B) + (D_{BA} - D_{AB})/2.$$

The asymmetry $D_{BA} - D_{AB}$ of a duplex link is expected to be quite small if similar equipment and the same transmission medium is utilized in the forward and reverse paths.

5. Independence of the Clock Error Measurement at Any Node from the Correction of Clock Error at Any Other Node (ICEM&C)

This is a feature which may be used to remove the effects of phase errors in clocks intermediate between the local node and the network master. This is accomplished by passing measured but uncorrected errors down the timing chains. Figure 5 shows how this is accomplished. When this feature is used in conjunction with double ended reference links and directed control, each node of the network is effectively tied to the network master. This feature cannot be effectively applied with the mutual control feature because in general there is no definite hierarchical path from the local node back to an ultimate reference.

6. Phase Reference Combining (PRC)

Figure 6 shows a model for the phase reference combining feature. This is a feature applicable for use in conjunction with DC, DE and ICEM&C which tends to reduce the effects of cumulative measurement errors in reference information arriving at nodes far removed from the network master. This is done by combining reference information received at a particular node over the parallel paths from the network master. It is designed to make good use of all the available reference information while carefully avoiding closed loops in the network. Two classes of two types of information are passed through the system when this feature is utilized. The two types of information are the actual measured values of the reference information and an estimate of the statistical variance in the measured value of the reference information. The statistical variance is obtained from design parameters of the network components. It is used to determine how to weight the measured reference information received over the various parallel paths. The two classes of information are called class 1 and class 2. At a particular node class 1 information of both types is obtained from neighbor nodes strictly above the local node in the network hierarchy. It is combined and used to discipline the local clock. The combined class 1 information is also passed to neighbor nodes not lower in the hierarchy than the local node. The class 2 information is obtained from neighbor nodes not lower in the network hierarchy than the local node. It is combined and passed to neighbor nodes lower in the hierarchy than the local node. The method used to combine the reference information is to weight it inversely proportional to the estimated variance.

This feature also provides for an elaborate monitoring capability.

7. Self Organizing (SO)

This feature is concerned with a scheme for distributed self organization of the network hierarchy. In a system that does not utilize the PRC feature the implementation is very similar to that described by Darwin and Prim [15]. Each node is assigned a rank with the network master being assigned a higher rank than any other node. The first alternate master is assigned a rank lower than the master but higher than any other node, etc. Each link is assigned a demerit number depending on the quality of the link. The object of the scheme is for each node to reference the highest ranking node through the highest quality path. It is shown in [15] that a network will actually do this in a stable manner when all rules are utilized.

For networks using the PRC feature as described by Stover [16] the link demerit information is essentially contained in the variance information so it is only necessary to know the number of intervening nodes between each node and the master reference, as well as the nodal rankings.

8. Master in Mutual System (MIM)

One of the disadvantages of the mutual control feature is that the network frequency may take random walks. This may be avoided by allowing one node to be the master, i.e., this master node does not reference any other node of the network. The result is that the long term average frequency becomes that of the master node.

9. Smoothing for Link Dropouts and Reference Switching (SLD&RSS)

The technique most commonly used to discipline the local clock is a second order phase locked loop. A step change in reference phase at the input to such a device results in a spike in output frequency whose peak amplitude is equal to $2\zeta\omega_n\Delta\phi$, where ζ and ω_n are parameters of the phase lock loop and $\Delta\phi$ is the reference phase change. This frequency spike can have a large peak amplitude and is quite undesirable. Fortunately, it can be avoided.

In a network using mutual control dropout smoothing consists of remembering the value of phase error relative to each reference immediately before the dropout. The remembered value is then applied through a decaying multiplier for a period of time sufficient to allow the remaining reference errors to slowly readjust, thereby avoiding the large spike in output frequency.

In a directed control network utilizing a second order loop of the integral plus proportional type the integrator voltage may be adjusted to exactly compensate for the difference between the old and new reference errors. Since the integrator has a long time constant the output frequency changes very slowly to compensate for the difference in phase references.

IV. THE SIMULATION MODEL

A computer simulation model was developed to help evaluate the ability of the set of features described in Section III to provide the set of desirable characteristics listed in Section II, though primarily for characteristics 1 through 6. This model consisted of the following:

- A nodal synchronizer
- A network topology
- A set of normal link and nodal variations
- A set of stress scenarios
- A set of feature combinations

1. The Nodal Synchronizer

Figure 7 shows a simplified functional block diagram of the nodal synchronizer. The weight and sum function is utilized in mutual control systems as well as directed systems utilizing the PRC feature although the implementation will be considerably different for the two cases. Under mutual control all available references are normally selected except that failed references must be deselected. Provisions are made for either a VCXO which is directly controlled by the loop filter output or an atomic clock whose output is indirectly controlled by means of an outboard phase stepper. The phase detector(s) was modeled as an extended range linear device. The loop filter was one of two types, having transfer function $\frac{a}{s+a}$ or $\frac{s+a}{s}$. The phase locked loop employing a loop filter of the first type results in a type 1 loop while one employing a loop filter of the second type results in a type 2 loop. The type 1 loop tracks a constant frequency offset (local clock's natural frequency offset from reference frequency) with a constant nonzero error signal out of the phase detector. The type 2 loop tracks a constant frequency offset with a constant zero error signal out of the phase detector. In order to avoid network instability problems only the type 1 loop was utilized with the mutual control feature. Either the type 1 or type 2 loop may be used with the directed control feature but the simulations were run almost exclusively with the type 2 loop

because its performance is clearly superior to that of the type 1 loop. Figure 8 shows a baseline model for the phase locked loop. For a detailed explanation of such devices see tutorial papers [17] and [18]. From this model the following transfer functions for the type 1 and type 2 loops can be derived:

$$\text{Type 1:}\quad \frac{\theta_o(s)}{\theta_i(s)} = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

where $\omega = \sqrt{ak_v}$ and $\zeta = \sqrt{a/4k_v}$

$$\text{Type 2:}\quad \frac{\theta_o(s)}{\theta_i(s)} = \frac{2\zeta\omega_n s+\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

where $\omega_n = \sqrt{ak_v}$ and $\zeta = \sqrt{k_v/4}$

Table 1 shows the loop parameters used in the simulation.

TABLE 1

Loop Parameters for Simulation

	ζ	ω_n (rad/s)
Type 2-DC-Quartz	4	5.6×10^{-5}
Type 2-DC-Cesium	2	1.12×10^{-5}
Type 1-DC&MC-All	1	1.52×10^{-3}

2. Network Topology

The topology of the network used in all simulations is shown in Figure 9. It was chosen to represent a skeleton of the DCS network in that link distances were chosen similar to those that will actually be utilized in the DCS network. The four gateway nodes were assumed to connect between the North American and European continents.

3. Normal Link and Nodal Variations

The normal link variations were modeled as sinusoidal length variations about the nominal length. All except the tropo links were modeled with a period of one day. Measurement jitter terms were added for each link and each node for some of the simulation runs. Initial frequency offset was assumed for each clock to give an initial transient. Quartz

clocks were given a linear drift in natural frequency. These variations are summarized in Table 2.

4. Stress Scenarios

The stress scenarios consisted of selected link failures, node failures, and clock frequency changes occurring at various times throughout the simulation runs. Each run represented an elapsed time of 300,000 seconds. The three sets of stress scenarios were called General, Low Level, and High Level. The General stress scenario was designed to measure how closely nodal phases and frequencies at various points of the network would hold together under a variety of disturbance events. The Low Level scenario was designed to measure any upward propagation in the network hierarchy. The High Level scenario was designed to measure propagation either upward, laterally, or downward. The General stress scenario is shown in Table 3.

5. Feature Combinations

A total of 16 feature combinations were selected for incorporation into a simulation run. These are shown in Table 4.

V. THE RESULTS AND CONCLUSIONS

Results:

Simulation runs were made for each of sixteen combinations of features and three stress scenarios for a total of 48 simulation runs. Each simulated run was for a duration of 300,000 seconds. Outputs consisted of plots of absolute frequency error and phase error relative to node 1 of Figure 9. Figures 10-15 show samples of these plots. The means and standard deviations of the frequency and phase errors were also obtained for each of the monitored nodes over the duration of the simulation run. The output plots and statistical data were then compared manually for those desirable characteristics which could be evaluated by this technique (primarily characteristics 1 through 6 of Section II). The other seven desirable characteristics were somewhat amenable to analysis in light of the broad understanding of the overall problem that was gained through the process of developing the simulation model. From this narrow simulation data and the broad problem understanding, relative rankings of the ability of each of the 16 feature combinations to provide each of the desirable characteristics was obtained. These rankings are shown in Table 5. In order to distinguish

Table 2

Normal Link and Nodal Variations

LINKS:

Link Type	1σ Measurement Jitter	Normal Link Variations
Microwave	10	$\Delta L = 10^{-5} x L_o x \sin(\omega_d t + \phi_r)$
Cable	10	$\Delta L = 3x10^{-6} x L_o x \sin(\omega_d t + \phi_r)$
Satellite	30	$\Delta L = 5.04x10^{-5} x L_o x \sin(\omega_d t + \phi_r)$
Tropo	20	$\Delta L = 10^{-8} x L_o x \sin(\omega_x t + \phi_r)$ $(10^{-4} x L_o x \sin(\omega_d t + \phi_r)$ $+2x10^{-5} L_o x \sin(8.7x10^{-4} t)$

where $\phi_r=0$, $\omega_d=7.3x10^{-5}$ rad/s, $\omega_x=3.5x10^{-3}$ rad/s and L_o=nominal link distance

NODES:

Node Numbers	1σ Measurement Jitter
1,2,13, & 14	10
3,4,5,6,7,8,15 & 16	20
9,10,11,12,17,18,19, & 20	30

Nodal variance and weighting factor calculation

$$\sigma^2 = \sigma_N^2 + \frac{1}{\sum_{i=1}^{n}\left(\frac{1}{\sigma_i^2+\sigma_{Li}^2}\right)} \; ; \quad \omega_i = \frac{\frac{1}{\sigma_i^2+\sigma_{Li}^2}}{\sum_{j=1}^{n}\left(\frac{1}{\sigma_j^2+\sigma_{Lj}^2}\right)}$$

where σ_i^2 or σ_j^2 is the input variance from reference i or j, σ_{Li}^2 is the variance associated with the link between the reference node and the local node and σ_N^2 is the nodal variance.

Table 3

General Stress Scenario

- Apply initial transient to all nodes
- Apply normal link delay variations to all links
- Apply clock drifts to all quartz clocks

50000 s	-	Node 6 clock starts ramp increase in frequency of $1.16 \times 10^{-14} \times$ fo per second and continues until 75000 seconds.
60000 s	-	Node 13's clock makes step decrease in natural frequency of $3 \times 10^{-11} \times$ fo.
75000 s	-	Node 6's clock begins ramp decrease of $1.16 \times 10^{-14} \times$ fo per second and continues until 100000 seconds.
100000 s	-	Node 13's clock makes step increase of $3 \times 10^{-11} \times$ fo in natural frequency.
150000 s	-	Link 6-5 fails. Node 5 free runs until 150300 seconds and then references Node 7 if self-reorganizing. Node 5 free runs until 160800 seconds if nonself-reorganizing..
200000 s	-	Link 1-2 fails. Node 2 free runs until 200300 seconds it references Node 6 if using self-reorganization, it free runs to 210800 seconds for nonself-reorganization.
250000 s	-	Node 13 fails. Node 16 free runs until 250300 seconds at which time it references Node 15 if using self-reorganization. It free runs to 271600 seconds if using non-self reorganization.

Monitor: Nodes 5, 6, 7, 2, 11, 13, 16, 17, 15, and 3

Links 7-5, 1-6, 1-7, 6-2, 5-11, 1-13, 13-16, 16-17, 14-15, and 2-3.

Table 4

Feature Combinations for Simulations

1. Directed control with Type 1 loop (mutual sync loop parameters).
2. Directed control with Type 2 loop.
3. Mutual control with equal weighting.
4. Mutual control with unequal weighting.
5. Mutual control with a master and equal weighting.
6. Mutual control with a master and unequal weighting.
7. Mutual control with dropout smoothing (and equal weighting).
8. Directed control with Type 2 loop and double-ended.
9. Mutual control with equal weighting and double-ended.
10. Mutual control with a master, unequal weighting, dropout smoothing, and double-ended.
11. Directed control with double-ended and independence of measurement and correction.
12. Directed control with double-ended, independence of measurement and correction, and phase reference combining.

SELF-ORGANIZING RUNS

13. Directed control with Type 2.
14. Directed control with double-ended.
15. Repeat Run No. 11.
16. Repeat Run No. 12.

Table 5 - Summary of Results

Characteristic / Feature Combination	1 Frequency Accuracy	2 High Level Clocks Not Disturbed by Perturbations at Lower Levels of Network	3 Clock Errors Do Not Harmfully Propagate to Other Nodes	4 Path Delay Variations and Dropouts Do Not Harmfully Propagate to Other Nodes	5 Slip Free	6 Survivable
1. DC-1-SE	13	1	3	--	--	--
2. DC-2-SE	12	1	3	10	2 (63)	9
3. MC-EW-SE	15	8	4	14	4 (100)*	14
4. MC-UEW-SE	16	5	4	15!	4 (100)*	12
5. MC-M-EW-SE	9	7	3	12	4 (100)*	11
6. MC-M-UEW-SE	10	6	3	13	4 (100)*	10
7. MC-DOS-EW-SE	14	4	4	11	4 (100)*	13
8. DC-2-DE	7	1	3	8	1 (13)	6
9. MC-EW-DE	8	3	4	9	3 (73)*	7
10. MC-M-UEW-DOS-DE	3	2	2	7	3 (73)*	3
11. DC-2-DE-ICEM&C	6	1	1†	6	1 (13)	5
12. DC-2-DE-ICEM&C-PRC	5	1	1†	5	1 (13)	4
13. DC-2-SE-SO	11	1	3	4	2 (63)	8
14. DC-2-DE-SO	4	1	2	3	1 (13)	3
15. DC-2-DE-ICEM&C-SO	2	1	1†	2	1 (13)	2
16. DC-2-DE-ICEM&C-PRC-SO	1	1	1†	1	1 (13)	1

In "Slip Free" column numbers in parenthesis indicate buffer sizes needed to avoid slips.

* Requires special provisions with quartz clock to cancel accumulated phase error every 100 days.

† Not harmful according to criterion.

! Potentially harmful according to definition.

Table 5 - Summary of Results
(Continued)

Characteristic / Feature Combination	7 Overhead Requirement (Bits/Sec)	8 Precise Time Available	9 Monitorability Incre-Mental	Overall	10 Federal Standard 1002	12 Inter-Operability	13 Selects New Master	11 Cost Effective Time (% Of 8080A)	Space (Bytes)
1. DC-1-SE	--	No	Fair	Good	Fair	Fair	No	0.77	395
2. DC-2-SE	--	No	Fair	Good	Fair	Fair	No	0.77	395
3. MC-EW-SE	--	No	Poor	Good	Poor	Poor	No	7.20	594
4. MC-UEW-SE	--	No	Poor	Good	Poor	Poor	No	7.20	594
5. MC-M-EW-SE	--	No	Poor	Good	Fair	Fair	No	7.20	594
6. MC-M-UEW-SE	--	No	Poor	Good	Fair	Fair	No	7.20	594
7. MC-DOS-EW-SE	--	No	Poor	Good	Poor	Poor	No	7.20	594
8. DC-2-DE	63	Yes	Good	Good	Fair	Fair	No	2.24	731
9. MC-EW-DE	63	No	Poor	Good	Poor	Poor	No	10.55	1045
10. MC-M-UEW-DOS-DE	63	Yes	Poor	Good	Good	Good	No	11.60	1164
11. DC-2-DE-ICEM&C	126	Yes	Very Good	Good	Good	Good	No	2.42	802
12. DC-2-DE-ICEM&C-PRC	171	Yes	Very Good	Good	Good	Good	No	10.65	1847
13. DC-2-SE-SO	76.5	No	Good	Good	Fair	Fair	Yes	1.46	702
14. DC-2-DE-SO	139.5	Yes	Very Good	Good	Good	Good	Yes	2.93	1009
15. DC-2-DE-ICEM&C-SO	202.5	Yes	Excellent	Very Good	Very Good	Very Good	Yes	3.11	1109
16. DC-2-DE-ICEM&C-PRC-SO	247.5	Yes	Excellent	Very Good	Very Good	Very Good	Yes	11.34	2154

harmful versus non-harmful propagation of disturbances, a criterion of harmful propagation was derived. This was primarily concerned with the probability that a disturbance (clock or link) would cause a bit slip in the input communications clock recovery loop (different from the output nodal clock) of a node down stream from the disturbance. This was predicated on our knowledge of the parameters of the clock recovery loops in existing DCS equipments. New equipments can be designed to handle disturbances of a given size and speed by selection of these parameters, albeit at a possible signal to noise ratio penality. The criterion is listed below:

$$\left|\frac{\Delta\omega}{\omega_o}\right| \geq 10^{-7} \quad \text{Definitely harmful.}$$

$$10^{-9} \leq \left|\frac{\Delta\omega}{\omega_o}\right| < 10^{-7} \quad \text{Potentially harmful and it causes SNR degradation.}$$

$$10^{-10} \leq \left|\frac{\Delta\omega}{\omega_o}\right| < 10^{-9} \quad \text{Unlikely to cause observable problems.}$$

$$\left|\frac{\Delta\omega}{\omega_o}\right| < 10^{-10} \quad \text{No effects on slip rate or SNR.}$$

Using this criterion against the desirable characteristic "clock errors do not harmfully propagate to other nodes" one clock error of 10^{-9} p-p was judged to be potentially harmful to some node of the network under each set of feature combinations except the ones indicated. In the "precise time availability" column the YES entries do not mean that precise time is automatically provided by this combination of features but rather indicate that the essential ingredients are provided to support an add-on unit to the nodal synchronizer which gives precise time. In the column titled "selects new master" the mutual control systems utilizing a master node (Runs 5,6, and 10) had no provisions for automatic selection of a new master and therefore were marked with NO. However, a real implementation of this combination of features would almost certainly require provisions for automatic selection of a new master.

Conclusions:

All features tend to do the things which they are logically designed to do. They all provide contributions to the attainment of subsets of the desirable characteristics.

Directed and mutual control tend to provide for a synchronized network with long term frequency averages at each node the same as or very close to the network frequency. Additionally, directed provides for long term zero average phase error and contains disturbances in the branch in which they occur and only propagates them downward.

Double-ended removes path delay variations.

Smoothing removes undesirable large frequency spikes due to step changes in reference phase.

Master in mutual system provides definite network frequency. Unequal weighting in mutual system tends to improve short term accuracy but can cause larger transients with link and nodal failures.

ICEM&C removes disturbances due to independent clock errors in a branch of nodes.

Phase reference combining is effective in combating measurement jitter and also decreases expected percentage of time that nodes may be without a reference during periods of stress.

- The simulations were performed with limited network size and connectivity and for very limited run time. With larger networks and connectivity and much longer run time the expected benefits from the additional features will be accentuated.

- Provisions for additional features (over and above mutual or directed control) does not seem excessive in that overhead data requirements are quite small and processor time and storage space is small in comparison with the capabilities of presently available microcomputers.

- The most striking of the additional features is the double-ended reference links.

- Although, according to the definition of harmful transient, most of the disturbances due to clock errors or path delay variations and dropouts experienced in the simulation were judged to be non-harmful these events were mostly isolated and their amplitudes were chosen to represent typical events. In a real stressed environment it is possible that several such events could occur closely enough together in time and at the correct points in the network to be harmful. The tabular summary

indicates the combinations of features least vulnerable to such threats.

- Precise time can be provided as an add-on feature to any scheme of control that has a master and utilizes double-ended links. The add-on does not affect the manner in which the nodal synchronizer controls the local clock's phase and frequency. Additional features may be used to improve the accuracy of disseminated time.

- Without a master the network frequency of the mutual control system may wander which makes interoperation difficult. Provision of a master would then dilute "claimed" survivability of this method of control. Overhead is also required to automatically select alternate masters.

- A disturbance occurring anywhere in the mutual control network propagates to all nodes of the network.

- In order to ensure network stability in a mutual control system limitations are placed on the type of nodal loop. The disadvantages are as follows:

  1. Type 1 loop tracks constant frequency offset with non-zero phase error.
  2. Type 1 loop tracks constantly drifting clock with non-zero frequency error and constantly increasing phase error.
  3. The above two characteristics tend to degrade the short term accuracy of the network.
  4. Characteristic 2. above indicates need for special provisions for drifting clocks, size buffer for lifetime operation, periodic adjustment of natural frequency, or other means.

- Error history at each node in mutual control system is a complex function involving network topology, reference weightings, stress events, and individual clock performance at all other nodes of the network. This makes it more difficult to devise a control strategy during intervals when a reference is not available. Thus, survivability is decreased. This complex history also lessens the utility of monitored parameters at each node.

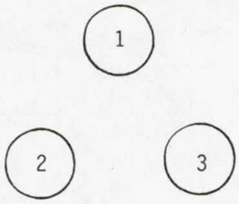


Fig. 1: Independent Clocks Network Model

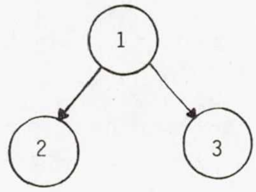


Fig. 2: Directed Control Network Model

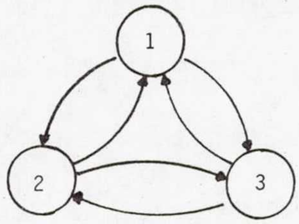
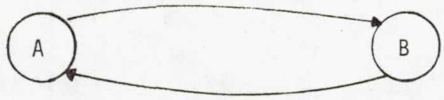


Fig. 3: Mutual Control Network Model

$$T_a, K_a = T_a - T_b + D_{ba}$$



$$T_b, K_b = T_b - T_a + D_{ab}$$

$$\frac{K_a - K_b}{2} = T_a - T_b + \frac{D_{ba} - D_{ab}}{2}$$

Fig. 4: Double Ended Reference Link Model

$$\Delta T_{AM} = T_M - T_A$$

$$\Delta T_{BM} = T_A - T_B + \Delta T_{AM}$$

$$\Delta T_{CM} = T_B - T_C + \Delta T_{BM}$$

Fig. 5: Independence of Clock Error Measurement and Correction Model

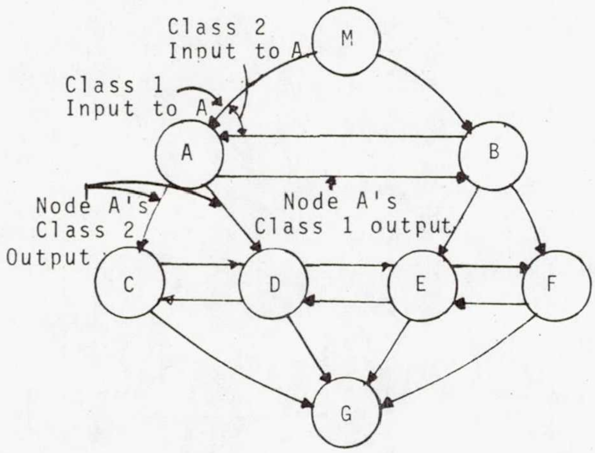


Fig. 6: Phase Reference Combining Model

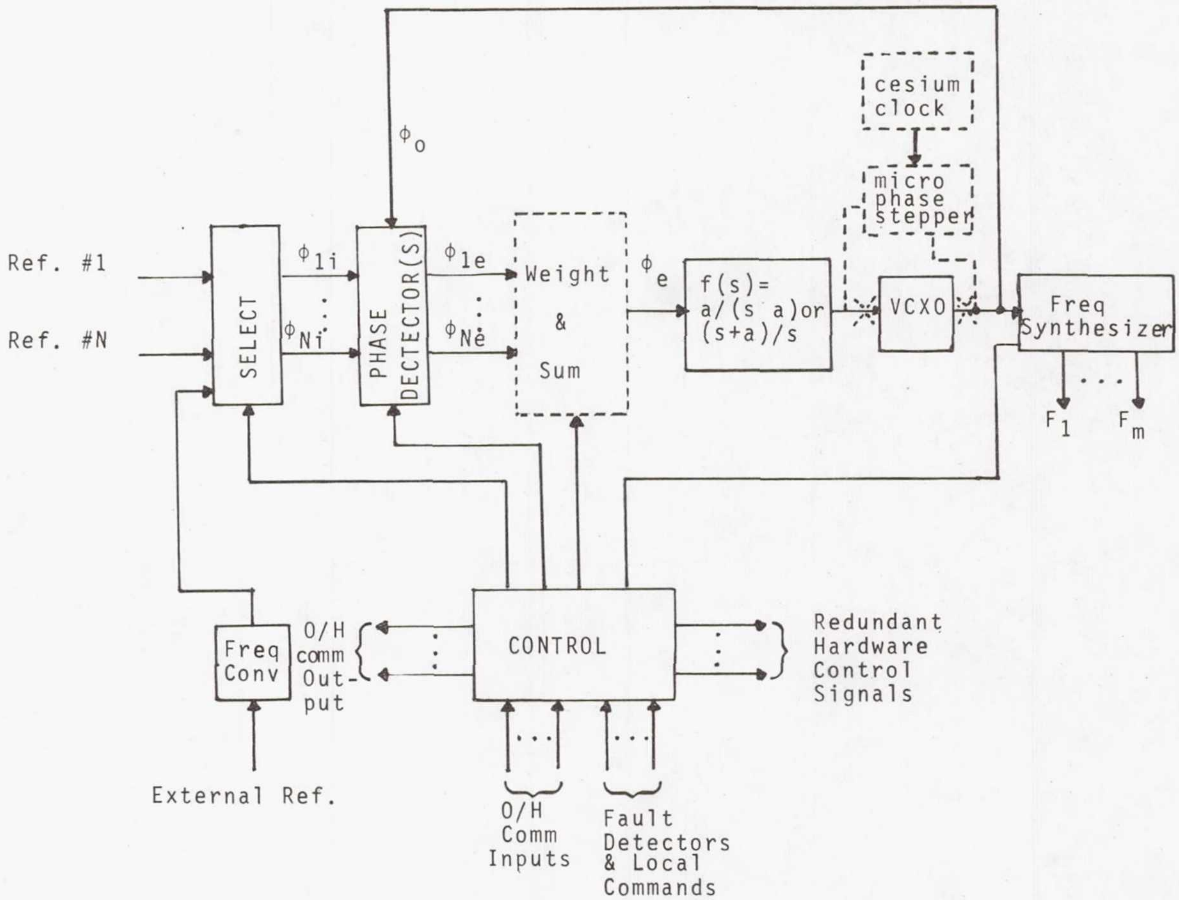


Figure 7: Functional Block Diagram for Nodal Synchronizer

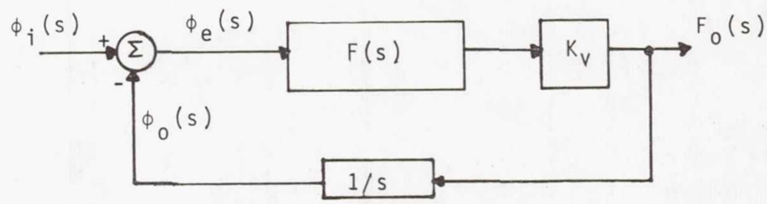


Figure 8: Baseline Phase Locked Loop Model

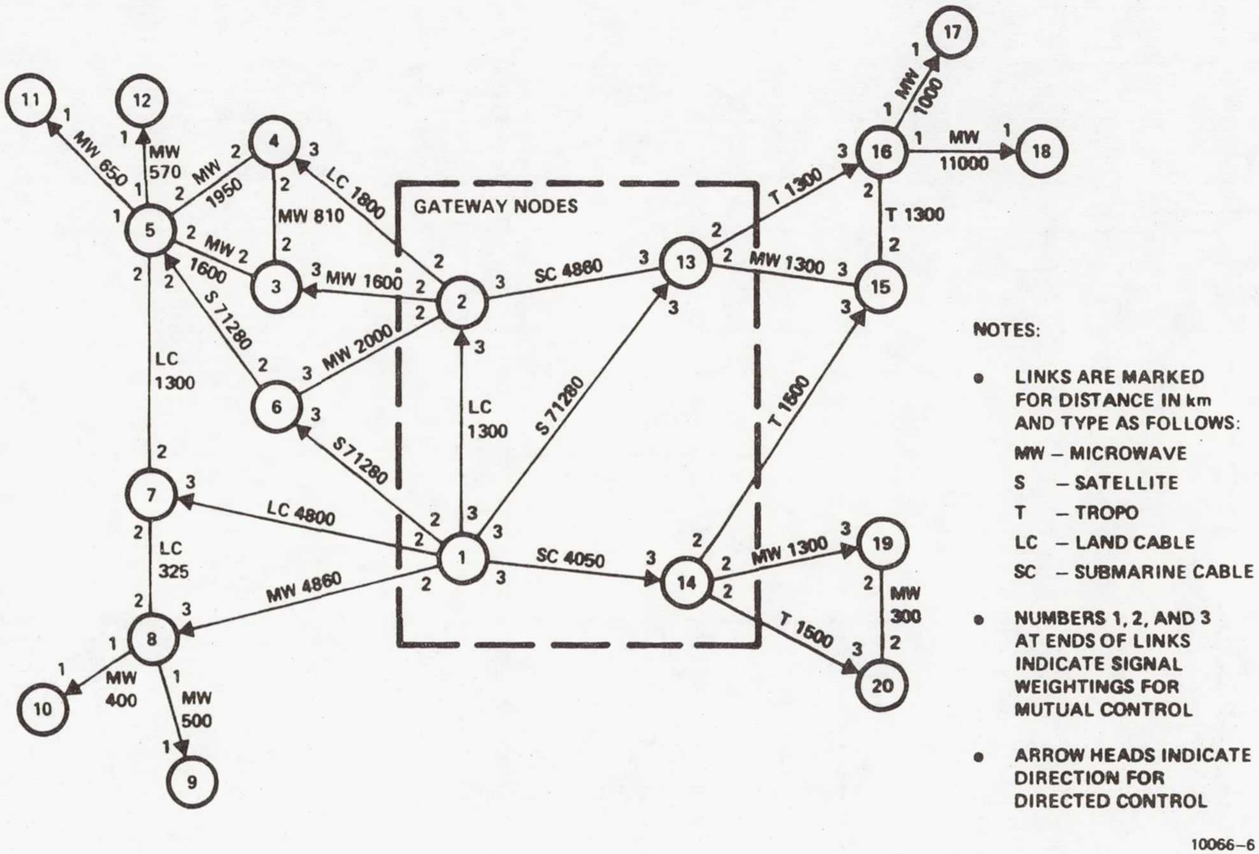


Figure 9: Network for Simulations

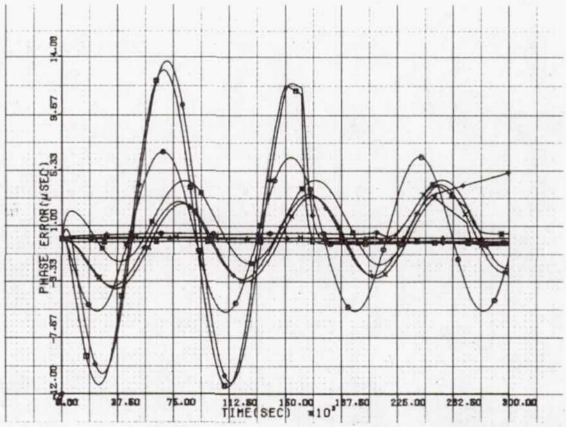


Figure 10 - Phase plot for directed control with type 2 loop and without drop-in smoothing and coating. General stress scenario.

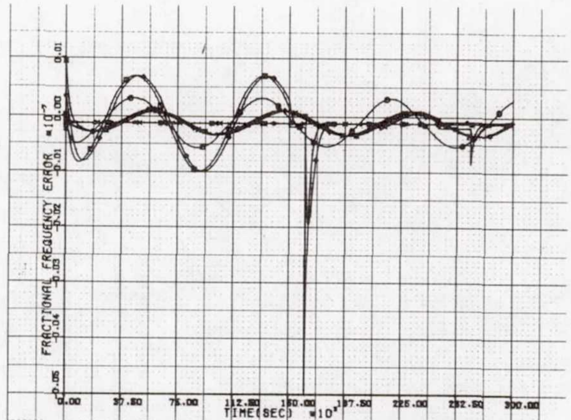


Figure 11 - Frequency plot for directed control with type 2 loop and without drop-in smoothing and coating. General stress scenario.

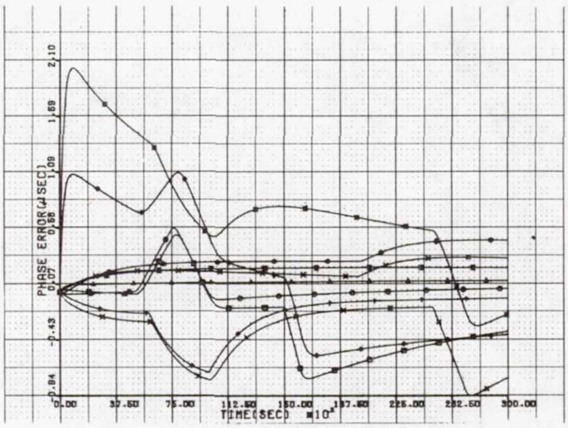


Figure 12 - Phase plot for directed control with type 2 loop and double-ended. General stress scenario.

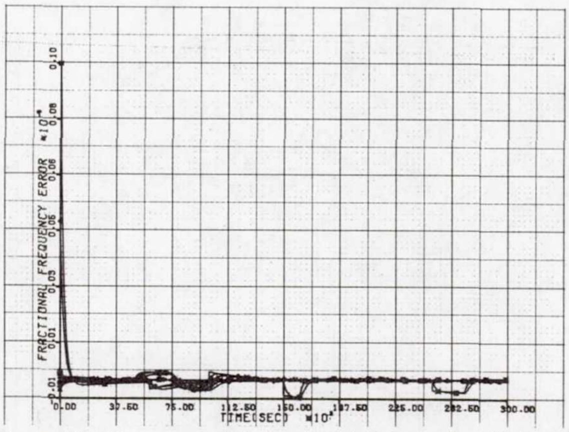


Figure 13 - Frequency plot for directed control with type 2 loop and double-ended. General stress scenario.

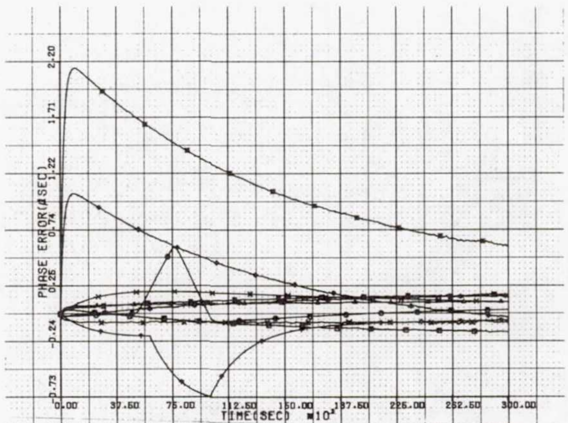


Figure 14 - Phase plot for directed control with double-ended, independence of measurement and correction, phase reference combining and self organizing. General stress scenario (with jitter).

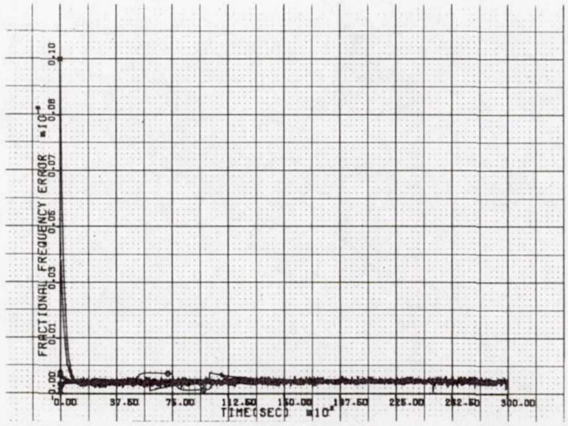


Figure 15 - Frequency plot for directed control with double-ended, independence of measurement and correction, phase reference combining and self organizing. General stress scenario (with jitter).

REFERENCES

[1] "Study of Alternative Techniques for Communication Network Timing/Synchronization," Harris ESD Final Report, DCA Contract DCA 100-76-C-0028, March 1977.

[2] J. G. Baart, S. Harting, and P. K. Verma, "Network Synchronization and Alarm Remoting in the Dataroute," IEEE Trans. on Communications, Nov. 1974, pp. 1873-1877.

[3] R. G. Dewitt, "Network Synchronization Plan for the Western Union All Digital Network," Telecommunications, July 1973, pp. 25-28.

[4] B. R. Saltzberg and H. M. Zydney, "Digital Data System: Network Synchronization," BSTJ, May-June 1975, pp. 879-892.

[5] M. B. Brilliant, "The Determination of Frequency in Systems of Mutually Synchronized Oscillators," BSTJ, Dec. 1966, pp. 1737-1748.

[6] M. B. Brilliant, "Dynamic Response of Systems of Mutually Synchronized Oscillators," BSTJ, Feb. 1967, pp. 319-356.

[7] A. Gersho and B. J. Karafin, "Mutual Synchronization of Geographically Separated Oscillators," BSTJ, Dec. 1966, pp. 1689-1704.

[8] M. Karnaugh, "A Model for the Organic Synchronization of Communications Systems," BSTJ, Dec. 1966, pp. 1705-1735.

[9] M. R. Miller, "Some Feasibility Studies of Synchronized Oscillator Systems for PCM Telephone Networks," Proc. IEE, London, 1969, pp. 1135-1143.

[10] J. P. Moreland, "Performance of a System of Mutually Synchronized Clocks," BSTJ, Sept. 1971, pp. 2449-2468.

[11] R. H. Bittel, W. B. Elsner, H. Helm, R. Mukundan, and D. A. Perreault, "Clock Synchronization through Discrete Control Correction," IEEE Trans. on Communications, June 1974, pp. 834-839.

[12] M. W. Williard, "Analysis of a System of Mutually Synchronized Oscillators," IEEE Trans. on Communications, Oct. 1970.

[13] M. W. Williard and H. R. Dean, "Dynamic Behavior of a System of Mutually Synchronized Oscillators," IEEE Trans. on Communications, Aug. 1971.

[14] A. C. Davies, "The Effects of Clock Drift Upon the Synchronization of Digital Communications Networks," IEEE Trans. on Communications, Nov. 1974.

[15] G. P. Darwin and R. C. Prim, U. S. Patent No. 2,986,723, "Synchronization of a System of Interconnected Units."

[16] H. A. Stover, "Improved Time Refererence Distribution for a Synchronous Digital Communications Network," Proc. of the Precise Time and Time Interval Planning Meeting, Nov. 1976.

[17] C. J. Byrne, "Properties and Design of the Phase-Controlled Oscillator with a Sawtooth Comparator," BSTJ, March 1962, pp. 559-602.

[18] Marc A. Rich, "Designing Phase Locked Oscillators for Synchronization," IEEE Trans. on Communications, July 1974, pp. 890-896.

QUESTIONS AND ANSWERS

MR. DAVID OWOLO, Western Union:

Most of the networks that both you and Ron touched on today used primarily terrestrial links for carrying either control or reference information to the various nodes, be they master or slave. What effects do satellite transmission links have on any of the systems you have discussed, and are there any special techniques that are more applicable to satellite linked networks than they are to terrestrial networks?

MR. WILLIARD:

I did touch on the fact in the first phase diagram I showed that there were satellite links assumed in three places in our network. The big phase variations, the 26 microseconds which I showed on that first phase plot, were essentially the result of the diurnal variations of the vertical drift of satellites in the directed control path to those nodes.

The use of double endedness can virtually eliminate those slow speed variations and not only the phase variations but the frequency variations also disappeared in that second set of graphs. The simple use of directed control will virtually eliminate the effects that satellites have upon the distribution and timing.

DR. HARRIS A. STOVER, Defense Communications Agency:

Now you have seen, I think, that the mutual is not the first choice probably, and whatever choice you make between directed control and mutual, it is generally beneficial to add the additional features after you have made that choice.

Page intentionally left blank

EFFECT OF VARIOUS FEATURES ON THE LIFE CYCLE COST OF THE TIMING/SYNCHRONIZATION SUBSYSTEM OF THE DCS DIGITAL COMMUNICATIONS NETWORK

D. B. Kimsey
Harris Corporation
Melbourne, FL

ABSTRACT

This paper reports the results of one task of a study, the purpose of which includes evaluating the advantages and disadvantages of implementing a set of optional timing/synchronization features in the future DCS digital communications network. The task reported on in this paper examines the effect on the life cycle cost of the timing subsystem when these optional features are included in various combinations. The features include mutual control, directed control, double-ended reference links, independence of clock error measurement and correction, phase reference combining, self-organization, smoothing for link and nodal dropouts, unequal reference weightings, and a master in a mutual control network.

An overall design of a microprocessor-based timing subsystem was formulated. The microprocessor (8080) implements the digital filter portion of a digital phase locked loop, as well as other control functions such as organization of the network through communication with processors at neighboring nodes. Other components of the subsystem include: extended-range linear phase detector, overhead interface, frequency standard, frequency or phase correcting apparatus, elastic buffers, and frequency synthesizers. A packaging scheme based on double-sided printed circuit cards was chosen. The particular packaging scheme was familiar to the author and is not claimed to be optimal for the application.

Seventeen timing subsystem configurations containing various combinations of the features were designed in enough detail to obtain reasonably reliable parts counts and estimates of the number of lines of code required for the 8080 routines. A hypothetical 200-node network (believed to be representative of a worldwide defense communications network) was used as the basis

for the life cycle cost analysis. For each of the seventeen configurations, the acquisition and life-cycle costs were computed based on 200 installations and a 20-year maintenance time.

The hardware acquisition and life-cycle costs were calculated using the RCA PRICE program (Programmed Review of Information for Costing and Evaluation). PRICE is a proven parametric cost estimating model which provides reliable estimates of system development, production, and maintenance costs. Software costs were calculated manually based on previous experience with development of 8080 software. Results of the pricing are presented in tabular form for each comparison.

INTRODUCTION

Various optional features may be included in the design of a synchronization subsystem to improve its performance. A companion paper [1] reports the performance effects of including such features. The features include mutual control, directed control, double-ended reference links, independence of clock error measurement and correction, phase reference combining, self-organization, smoothing for link and nodal dropouts, unequal reference weightings, and a master in a mutual control network.

The hardware and software implementations of the Timing Subsystem features are addressed in this paper. Seventeen Timing Subsystem configurations possessing these features in all practical combinations were designed and their life-cycle costs were determined. The costs of these configurations may be used in conjunction with the performance evaluations in [1] to evaluate performance versus cost trade-offs.

The Timing Subsystem consists of all components not included in the radios, modems, multiplexers, data terminals, etc., which are necessary to provide timing for and to synchronize data transfer between said equipments. The broad categories of these components are: local reference, buffers, timing generation and distribution, and optional disciplining circuitry.

Conceptual Design

Figure 1 shows the assumed multiplexer hierarchy used for the design analysis. It should be noted that a new generation of equipments may be desirable for use in the future synchronous DCS. The Timing Subsystem was designed around the present equipment specifications as much as practical, and the resulting Timing Subsystem costs may be somewhat higher as a result.

Figure 2 is a generalized block diagram of the Timing Subsystem. The system depicted here contains the maximum hardware and can implement all of the optional features. The detailed implementation of each of these components is described below.

The elastic buffer absorbs rate variations between the received timing and the local clock. An elastic buffer is placed between each of the two channels of the digital radio and the two corresponding TD-1193 Demultiplexers. (Only one is shown for brevity.) These buffers could be placed lower in the demultiplexer hierarchy, resulting in a larger number of smaller, slower buffers. An analysis of optimum buffer placement was not included in this effort; they are assumed to be between the radio and highest-level demultiplexer for ease of analysis.

Approaches which discipline a clock based on time of arrival of link timing signals require a sync extractor and phase detector. The sync extractor searches the mission bit stream for the TD-1193 sync patterns. The phase detector measures the time differences between the predicted and actual arrival times of these sync patterns. The time differences (phase errors) are inputs to the phase-locked loop which serves as the local time reference.

The processor inputs these phase errors and calculates the difference equations which implement the digital loop filter. The output samples from this filter discipline the nodal clock. They may control a VCXO (voltage controlled crystal oscillator) via a D/A (digital-to-analog) converter, or may discipline a crystal or atomic standard by shifting its phase via an outboard phase shifter. The nodal clock drives the frequency synthesizers which produce all phase-related timing required by the multiplexers, radios, modems, etc., and all demultiplexers which are lower in the hierarchy than the elastic buffers.

Timing approaches which employ overhead information require an overhead interface which allows information to be exchanged between processors on opposite ends of the link. It is assumed that this information is transmitted through a low-speed channel (300 b/s) of the orderwire multiplexer (LSTDM).

Each of the features described below was analyzed to determine its hardware and software requirements.

Implementation of Features

Directed control requires at least one sync extractor and phase detector in the system. For single-ended systems, the input to this circuitry is multiplexed so that it may connect to any one of the incoming links. This feature also requires a loop filter (processor) and means for controlling the local clock.

Double-endedness under Directed Control requires that the phase error measurement be made on each end of the link and that the master transmit its measured phase error to the slave via the overhead channel. The local node must contain a sync extractor and phase detector for the link to which it is slaved, and a sync extractor and phase detector for each neighboring node which is slaved to the local node. Further, the local node must contain an overhead receiver for the link to which it is slaved, and an overhead transmitter for each neighboring node which is slaved to the local node.

Double-Endedness under Mutual Control requires an overhead transmitter and receiver for each connecting link.

This feature also requires an additional computation (subtraction and division by 2) to be performed by the processor. Under Mutual Control, this computation is required for each connecting link.

Independence of clock error measurement and correction is only implemented after implementing both Directed Control and Double-Endedness. No additional hardware is necessary. The Master transmits its measured-but-uncorrected error (relative to the ultimate master) to the slave along with the measured phase error used for double-endedness. The slave subtracts this number from the result of its double-endedness calculation. Thus, only a minute amount of additional software is required. It should be noted that the Master node could combine these two numbers and transmit them as one number, thus requiring no additional bandwidth.

Phase reference combining is implemented only after implementing Directed Control, Double-Endedness, and Independence of Measurement and Correction. The local node derives timing from all nodes not lower in the hierarchy, and therefore must contain a sync Extractor and phase Detector for each of these links. Further, all nodes not higher in the hierarchy derive their timing in part from the local node. Since a double-end measurement is used, a sync extractor and phase detector are required on these links. Thus, a sync extractor and phase detector are required for each link.

The usual overhead transmission required for double-endedness, plus the measured-but-uncorrected phase error and a variance estimate, must be received from all nodes not lower in the hierarchy. The same three information types must be transmitted to all nodes not higher in the hierarchy. Thus overhead receivers are required on all links connecting to nodes not lower in the hierarchy, and overhead transmitters are required for all links connecting to nodes not higher in the hierarchy. For example, if the local node connected to three nodes higher in the hierarchy, three nodes of the same level, and four nodes lower in the hierarchy, it would require six receivers, and seven transmitters.

Additional software is required to implement Phase Reference Combining.

Rule 11, as described by Stover [2] requires the three data types to be exchanged in both directions on all links. This additional information is used for diagnostic purposes only, and is not required for implementing Phase Reference Combining. If Rule 11 is implemented, then overhead receivers and transmitters are required on all links, and additional software is required.

Self-organization requires overhead information to be passed in both directions over all links, i.e., the processor at the local node has full duplex communication with the processors at all neighboring nodes. Thus, an overhead receiver and transmitter is required on each link.

Additional bandwidth and processing are also required. For a Directed Control System not containing Phase Reference Combining, each node selects the best link to serve as its reference based on three data types: nodal rank, distance from master, and link demerit. This scheme is used in the earlier Time Reference Distribution [3] and adheres to rules similar to those of Darwin and Prim [4]. If Phase Reference Combining is included, the self-organization feature is implemented differently. The local node does not have to select the best link for its reference since it is always deriving timing from all neighbors not lower in the hierarchy. However, it must know at all times which of its neighbors are higher, lower, or on an equal level within the hierarchy. Two information types, INFO 1 and INFO 2, are employed for this determination.

Mutual control requires a sync extractor and phase detector on each incoming link. In addition to the loop filter software (which is of equal complexity with that required for Directed Control) the Mutual Control feature requires a weighted average of the phase errors derived from the individual links.

HARDWARE/SOFTWARE REALIZATIONS

The various components shown in Figure 2 were designed in enough detail to determine a parts list plus space and power requirements. A standard packaging scheme widely used at Harris ESD was selected as the basis for design. No claim is made that this packaging scheme is optimal for this particular application. This scheme employs 4.5 inches x 5.25 inches double-sided printed circuit (PC) cards which plug into motherboards via 80-pin connectors. A drawer was designed to meet the EMI requirements which have been specified for such equipments as the digital radio and various multiplexers. Two motherboards (each holding up to 23 cards) can be mounted horizontally in the front of the drawer, with enough room in the rear for power supplies.

Two drawers were necessary to house the circuitry for most configurations. It was deemed desirable to separate the components which would be common to all approaches from those which would be configuration dependent. Components common to all approaches are the frequency synthesizers, distribution amplifiers, and elastic buffers. These components were packaged in a single drawer herein called the Basic Drawer, which is constant across all configurations. This drawer was included in the cost figures to keep the cost of the various features in perspective; its cost may be easily factored out to more closely compare the costs of the various approaches used in the second drawer (herein called the Disciplining Drawer).

With the exception of the frequency synthesizers, which contain some ECL (Emiter-Coupled Logic), the designs incorporate the more economical and less power consuming low-power Schottky TTL (Transistor-Transistor Logic). Both drawers contain power supplies and motherboard wiring sufficient to support the PC cards required for seven terminating links (the assumed maximum). It was assumed, however, that the <u>average</u> node would only interconnect with four other nodes. Thus the costing was accomplished assuming a main frame capable of supporting seven links, but populated with PC cards to support four links. The following paragraphs describe the implementations of the individual components.

Phase Detector

Figure 3 depicts the sync extractor and phase detector circuitry. The clock output from the digital radio and the local 10 MHz reference are divided down to a common 8 kHz where the phase comparisons are made.

Phase difference is measured by counting cycles of the 10 MHz reference between the rising edge of the 8 kHz wave derived from the link timing, and the falling edge of the 8 kHz wave derived from the local 10 MHz reference as shown by the diagonal arrows in Figure 3.

The detection of synchronization patterns occurring at some submultiple of 8 kHz (depending on the selected output rate of the TC-1193) results in pulses from the sync detector synchronizing the countdown chain to the received framing. A countdown chain from the local reference controls the time of departure of the local TD-1193's frames by synchronizing both the TD-1193 <u>and</u> the transmit portion of the digital radio to the local clock.

Accurate phase measurements are obtained by averaging the counts obtained in several successive measurements. Results from [1] indicate that a phase measurement needs to be read out no more often than 1.5 times per second, or every 0.667 ms. This would allow in excess of 5,000 successive measurements to be averaged. Averaging is simply accomplished by allowing the counts to accumulate in the accumulator

counter until readout time, and dividing by the appropriate constant. At zero phase error, an average of 625 counts will be accumulated for each measurement. If the interval counter is configured to accumulate 5,000 or more such measurements, the resulting granularity of measurement is approximately 2 ns which is better than the asymmetry of the link and the associated equipments and is thus more than an adequate measurement granularity. This phase detector has an extended range of ±62.5 μs. The sync extractor requires one PC card, and the phase detector requires two cards.

Loop Filter and Overhead Processor

In order to implement a filter with a time constant on the order of several days, a digital (as opposed to analog) filter is a necessity. An 8080 microprocessor was chosen to implement this filter. The processor card designed is a self-contained computer including 1K bytes of Programmable Read-Only Memory (PROM) and 1K bytes of Random Access Memory (RAM). This single card is sufficient to perform the loop filter function. An analysis of software requirements is presented in [1].

Computations and bookkeeping required for implementing the overhead functions can be handled by the same processor used for the loop filter. Memory requirements for these features are presented in [1]. When more than 1K bytes of program storage is needed, a 4K byte PROM card is added to the system.

Technology advances in this area can quickly obsolete the results of cost/performance studies. As of this writing, 2K byte and 4K byte PROM IC's are becoming available which can replace the 1K PROM on the processor card. Intel has recently announced an 8K byte mask programmable read-only memory (ROM).

Local Reference

The assumed crystal reference is a 5 MHz oscillator having a drift of 1×10^{-10} per day. It is a self-contained, rack-mountable unit with its own power supply and stand-by battery system. A voltage input of ±5 volts will deviate the 5 MHz output by $\pm 2 \times 10^{-8}$. The short term stability is 1×10^{-11}. Many manufacturers, including Hewlett-Packard, Austron, Vectron, and Frequency and Time Systems, Inc., offer very similar references of this type with approximately $3,000 price tags. A frequency doubler is used to obtain the 10 MHz.

The crystal reference is disciplined with the output voltage of a D/A converter. A frequency resolution of at least 10^{-11} and a range of 4×10^{-8} require 4,000 quantization steps resulting in a 12-bit requirement for the D/A. This range allows the oscillator to drift for

200 days before the center frequency must be mechanically adjusted. This is the approach which was costed. Use of a $\pm 3 \times 10^{-7}$ adjustment range with a 16-bit D/A would permit the same resolution with a reset interval of 8 years, which exceeds the 5 year MTBF of the reference. However, this arrangement would have a sensitivity of 150 μvolt per quantization step on the voltage input. It would be extremely difficult to prevent noise pickup of this amplitude from modulating the reference.

The assumed Cesium Clock is a Hewlett-Packard Model 5061A with the standby power supply and high performance tube options; total price is \$22,750. This is a self-contained reference having an accuracy of $\pm 7 \times 10^{-12}$. The reference is disciplined by shifting its phase with an Austron 2055A Phase Microstepper costing \$3,550. The Rubidium standard is a Hewlett-Packard Model 5065A with the standby power supply option; total cost is \$8,575. This is a self-contained reference having a drift of $\pm 1 \times 10^{-11}$ per month. For disciplined approaches, its phase is shifted with the Phase Microstepper.

Overhead Interface

The Transmitter and Receiver portions of the Overhead Interface are shown in Figure 4. A total of 55 bits of information is required for implementing all optional features. Triplicating this figure results in 165. Adding 35 bits for framing and time of day information results in a total of 200 bits. Transmitting this information 1.5 times per second results in a 300 b/s data stream which may be transmitted via a standard 300 b/s channel of the LSTDM and the Digital Radio Orderwire. This information would only occupy 0.16 percent of the 192K bandwidth allocated for the orderwire.

The Transmitter accepts the information words 8 bits at a time from the 8080, serially transmits each word three times, and generates periodic framing information to allow separation of data on the receiving end. The Transmitter occupies one PC card.

The Receiver converts the serial stream to parallel, accumulates three successive words, votes to correct errors, and presents the parallel data to the 8080. A sync correlator detects the presence of the sync pattern and initializes the sequencer to pick out the words at the proper time. The Receiver occupies two PC cards.

Frequency Synthesizers and Distribution Amplifiers

The function of the frequency distribution system is to generate phase-related rates to clock all equipments which are to operate synchronously with the local clock. Such equipments may include all devices in the transmit hierarchy from data terminal equipment to the digital Radio, and all devices lower in the receive hierarchy than the elastic buffers.

built less expensively if the MUX and DEMUX do not have to operate from independent timing.

The 12.928 MHz data rate from the Digital Radio poses no difficult design problem for the Elastic Buffer. For an independent clock approach employing Cesium clocks and a 24-hour buffer reset interval, the required buffer size (at 12.928 MHz) would be 46 bits. For a Rubidium clock with a 10^{-11} per month drift, 6 month recalibration interval, and 24-hour buffer reset interval, the required buffer size is 270 bits. If disciplined nodes containing crystal clocks can be controlled to within 10 μs of the Master (within 20 μs of each other) then $\pm 259 = 518$ bits are required during normal operation. A node with a good crystal clock will drift an additional ± 56 bits during the first 24 hours after being severed from the network. Provided this is enough time to get the link back up, a total of 630 bits are required based on the above assumptions.

Figure 6 is a block diagram of a 1024-bit buffer designed to operate at 12.928 MHz. For relatively small buffers commercially available FIFO's (First-In-First-Out Memories) are perhaps the best approach. These IC's contain the control circuitry for moving data bits forward whenever one is extracted from the output. The FIFO in this configuration need only operate at 1.616 MHz. Bits from the 12.928 MHz stream are serially accumulated and stored broadsize (eight at a time) at 1/8 the original rate. The FIFO array can be implemented with either four 64 x 4 IC's or with four 32 x 8 IC's. Both types are available which can operate at the indicated rates. Handshaking signals are available to allow the IC's to be cascaded. The additional circuitry in Figure 6 is required to initialize the buffer (inhibit output clock until it half fills) and to monitor overflow and underflow. Two of these buffers occupy one PC card.

LIFE CYCLE COST ANALYSIS

It was desired to determine the cost of adding each one of the optional features to a basic timing approach. However, it is not possible to examine the cost of each feature individually. Rather, due to commonality of required components, it is more feasible to cost all practical configurations which include the features in various combinations. The list of configurations simulated in [1] was chosen for costing, so that performance versus cost trade-offs may be made. It became readily apparent that many of the configurations were practically identical with respect to cost. Estimating cost differences between such configurations was beyond the precision of the methods used for costing, and such configurations were combined into one. This resulted in reducing the 16 configurations presented in [1] into ten slightly more general configurations. The independent clock approach was added for completeness, resulting in eleven configurations which were costed.

DCEC specification R220-77-2 describes such a system which operates from the 1 MHz outputs of the AN/GSQ-183 Loran Receiver. Table 1 is a list of rates from that specification along with the number of required outputs of each rate. A balanced low-level driver (MIL-188-114) must be used for each output. This system is being procured for use in the interim communication network (DCS II), and might possibly be useable for the future DCS.

The assumed design generates the rates of Table 1 from the output of the 10 MHz local reference. Each of these rates, as well as the 10 MHz, is a multiple of 8,000 b/s. The phase of these rates should be such that if each one (including the 10 MHz input) is divided down to 8,000 b/s, then the 8,000 b/s waveform from each countdown chain should be in phase. The rising edges of this 8,000 b/s waveform (or a submultiple thereof) should be used to initiate the frame departure in the TD-1193 and Digital Radio.

Figure 5 depicts the design of the frequency distribution system. Each family of rates is generated by a VCXO and countdown chain. The VCXO's are locked via a broadband phase-locked loop to the 10 MHz reference. Four PC cards were required to implement these phase-locked loops.

The line drivers were implemented with commercial devices having voltage swings similar to those of MIL-188-114 balanced drivers. Devices which conform rigidly to the MIL-188-114 specification are not commercially available and must be special-made. Some companies (such as Sperry) have developed hybrid circuits which can be produced on special order. Typical cost is $80 each. Table 1 implies that 220 such devices are required. This number of drivers required 18 PC cards in the assumed design.

Elastic Buffer

The Elastic Buffer absorbs rate variations between received data and the local clock. They may be placed anywhere in the Demultiplexer hierarchy as long as the point where timing is derived is higher in the hierarchy than the buffers. Demultiplexers higher in the hierarchy than the buffers derive their timing from the associated incoming links, and those lower in the hierarchy receive timing from the nodal clock.

Since all demultiplexers lower in the hierarchy than the buffers are synchronous with each other and with the multiplexers, channel outputs from such demultiplexers may be routed to channel inputs on any multiplexer for retransmission on another link. This is a strong argument for placing the buffers as high in the hierarchy as possible. Placing them between the radio and the TD-1193 makes tandeming at any level possible, including routing one of the radio channels directly to another radio for transmission. A multiplexer/demultiplexer set can be

These configurations are listed in Table 2 using the same configuration numbers as in [1] for clarity.

Software and Hardware Requirements

Software costs were computed based on number of lines of code. A line of 8080 code expands into one, two, or three bytes (or no bytes if a comment statement) with the average being slightly greater than 2 bytes per line. Twenty percent was added to the requirements presented in [1] to accommodate diagnostic software. Table 3 lists the software requirements in terms of bytes and lines of code. For configurations requiring more than 1K bytes of memory, the PROM card is added to the system.

As previously stated, the synthesizers, distribution amplifiers, and buffers were placed in a separate drawer (the Basic Drawer). This drawer is included in all configurations. Table 4 is a breakdown of its contents. A fractional motherboard indicates only part of it is wired. The motherboard wiring and power supplies will support the maximum (Max/Box) number of PC cards expected. Costs were based on the average number of PC cards.

All configurations, except Independent Clocks, contain a second drawer (the Disciplining Drawer). Table 5 is a breakdown of the components contained in this drawer for all configurations. The motherboards and power supplies vary in size with the configurations, and in each case support the maximum number of links. Table 6 is a breakdown of the PC cards whose quantities vary between configurations. Costs are based on the average number. The choices of numbers in some cases are rather subjective and are based on assumptions of the number of neighbors higher, equal, or lower in the hierarchy.

Costing Methodology

The Software and Hardware costing were performed separately. The RCA PRICE program was used for Hardware costs. At the time of the costing, the PRICE Software model was not available. The Life-Cycle Cost of a system is divided into three parts: Development, Production and Maintenance for the life of the equipment.

Development costs include equipment design and construction of prototypes. These costs are nonrecurring; i.e., they are independent of the quantity of systems to be built. Production costs include tooling up for production, material and labor for building each system, and labor for testing finished systems. Tooling includes procuring or building special equipment used for fabricating and testing the systems. Production costs are proportional to the quantity of systems, but the relation is not linear. Due to a "learning curve," the cost of

production on per system basis decreases with the number of systems. Much of the tooling is up front; however, retooling generally is necessary due to wear and breakage. Maintenance costs include test and repair labor costs, transportation, supply management, and purchase of piece parts.

Acquisition costs simply consist of development costs plus the costs to produce the desired number of systems. When considering Life-Cycle Costs, the Production Costs are modified to include production of spares and production (or purchase) of test equipment to support the system in the field for a specified number of years.

Generally, software costs are considered to be nonrecurring. The costs simply consist of writing the code. If the software is the same in all systems, the software costs are independent of the number of systems. A reasonable cost for developing code of this type is $15 per line. For the system under consideration, the programs must be "burned" into the PROM IC's for each 8080 or PROM card. This process is mechanized and may be considered to be part of the fabrication process. These costs are included in the hardware production costs by assuming the complexities of these cards to be slightly higher than otherwise would have been assumed. Thus no additional production costs for software were assumed.

Software maintenance is a euphemism invented in recent years to describe the costs of continually rewriting programs which were not written properly in the first place. Many of these "errors" result from simply not anticipating every possible situation with which the software might have to deal. Even after "thorough" testing, residual errors may become apparent only after very long periods of operation. As a result, some software support may be required for the life of the system. A figure of $5 per line per year has been determined as a typical figure for such support.

The Life-Cycle Costs for the hardware configurations were computed using the RCA PRICE (Programmed Review of Information for Costing and Evaluation). There are two programs involved: PRICE 83B computes Acquisition Costs, and Price L1 modifies the production costs and adds in maintenance to complete the Life-Cycle Costs. PRICE is a proven parametric cost estimating model which provides reliable estimates of system acquisition costs (development and production). The PRICE 83B program generates design to unit production cost, based upon variations in designs, performance, schedules, reliability, economic escalations, etc. The price inputs are primarily physical characteristics of the design concept. These include weight, volume, manufacturing complexity, platform, quantity, development schedule and production schedule. The outputs feature recurring and nonrecurring costs for development and production as well as a unit production cost value for each entry.

PRICE 83B also develops inputs for the PRICE L1 model for Life-Cycle cost (LCC).

The Life-Cycle Costs of a system of 200 nodes were computed based on a system lifetime of 20 years. The software costs were computed by the simple formulas stated above. Table 7 is a breakdown of the software costs for the various configurations.

Hardware costs were computed by a very detailed process. The hardware was described in detail to PRICE 83B which computed Acquisition costs. This process was performed by personnel who are very experienced with the operation of the PRICE Model. These descriptions were performed at the PC card level and were checked for reasonability. The PC cards were combined in the various combinations with purchase items (items such as power supplies and references for which catalog prices were used) to form the different configurations. Another output of PRICE 83B is the LC file which includes unit costs, computed MTBF and MTTR values, and other pertinent factors used by PRICE L1 to compute the Life-Cycle Costs. Editing this LC file gives the user the opportunity to provide all the information he can about the system. For example, the predicted MTBF's of purchase items were overridden at this point with actual values from the manufacturer. MTBF's for designed equipment were also checked and altered if unreasonable.

The L1 Model was then exercised on the LC file using the force structure of Fig. 7. The support philosophy adheres to DoD Directive 4151-16 which states that there shall be three echelons of support. Simple repair is performed at the organizational shop, and consists of fault isolation to an LRU (Line Replaceable Unit), replacement with a spare and shipment to a higher level for repair. Generally, repair to piece part is performed at an Intermediate shop if not too complex, and at a Depot otherwise. Since most DCS nodes will probably be located at major military installations, the Intermediate shop is considered to be a general repair shop local to the base. It was assumed that 25% of the nodes would be remote, so that a total of 150 Intermediate shops were used. A Depot was assumed for each of the three services who will support DCS.

CONCLUSIONS

Table 8 is a breakdown of the components costs for Configuration 16 (using crystal clocks). Table 9 gives the Life-Cycle Costs of the various configurations. These costs are based on crystal clocks for the disciplined approaches. Configuration 16 is repeated for Cesium and Rubidium. Configuration 17 is also given for both Cesium and Rubidium but not Crystal.

The prices obtained in Table 9 may be compared with the simulation

results for cost versus performance trade-offs. Comparing the costs of disciplined approaches using crystal clocks, they are all very close (about 25 percent total variation). A surprising result was that the Life Cycle Cost of most disciplined approaches using crystal clocks came out slightly higher than that of Independent clocks using Cesium clocks.

REFERENCES

1. D. B. Bradley, J. B. Cain, II, and M. W. Williard, "An Evaluation of Optimal Timing/Synchronization Features to Support Selection of an Optimum Design for the DCS Digital Communications Network," Proceedings of the Tenth Annual Precise Time and Time Interval (PTTI) Planning Meeting, Nov. 1978.

2. H. A. Stover, "Improved Time Reference Distribution for a Synchronous Digital Communications Network," Proceedings of the Precise Time and Time Interval Planning Meeting, Nov. 1976.

3. H. A. Stover, "A Time Reference Distribution Concept for a Time Division Communication Network," Proceedings of the Fifth Precise Time and Time Interval Planning Meeting, Dec. 1973.

4. G. P. Darwin and R. C. Prim, U.S. Patent No. 2,986,723, "Synchronization of a System of Interconnected Units."

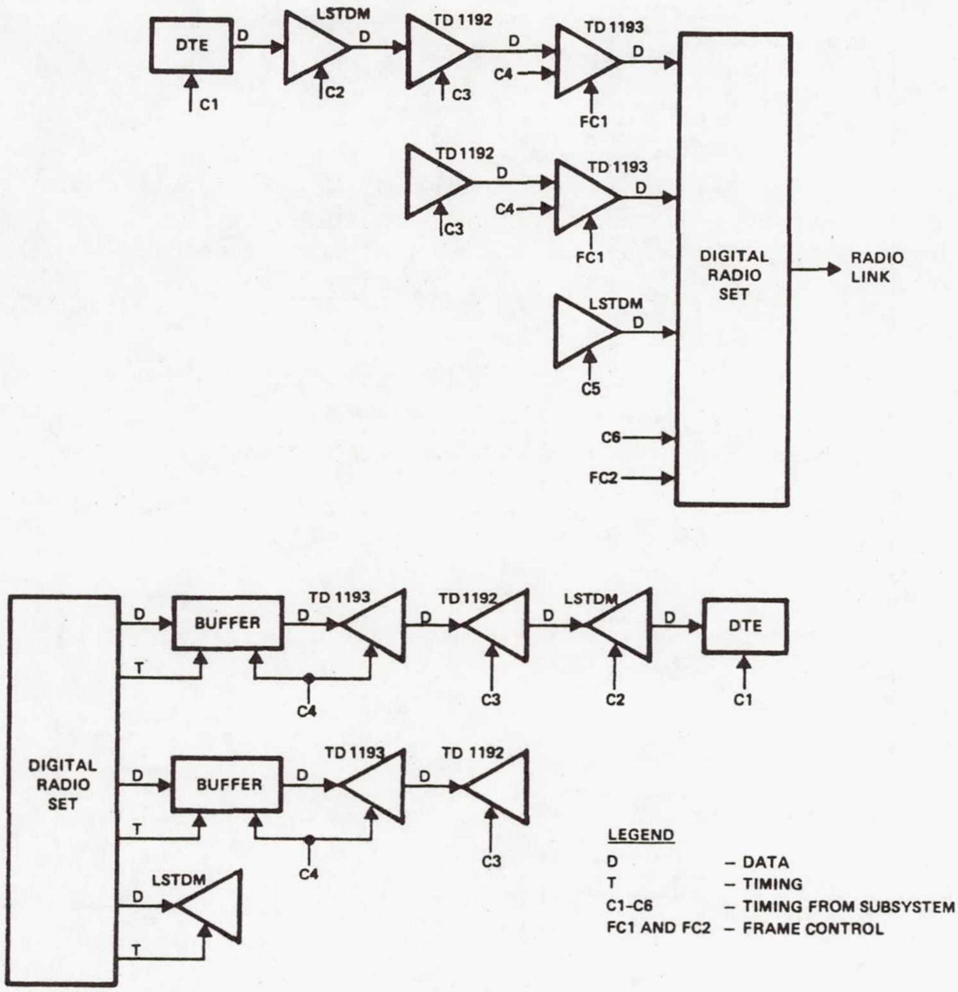


Figure 1. Transmit and Receiver Configurations

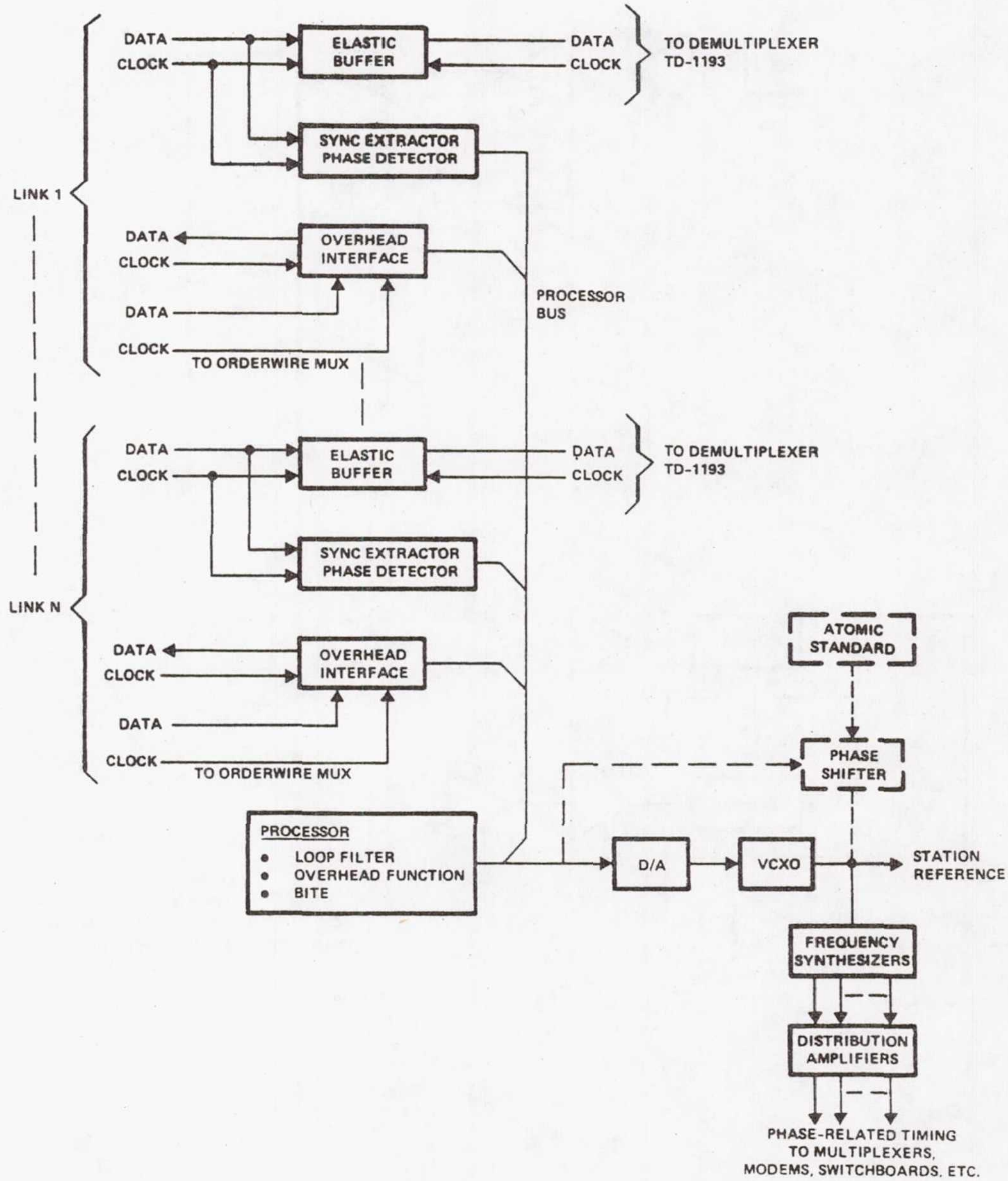


Figure 2. Timing Subsystem

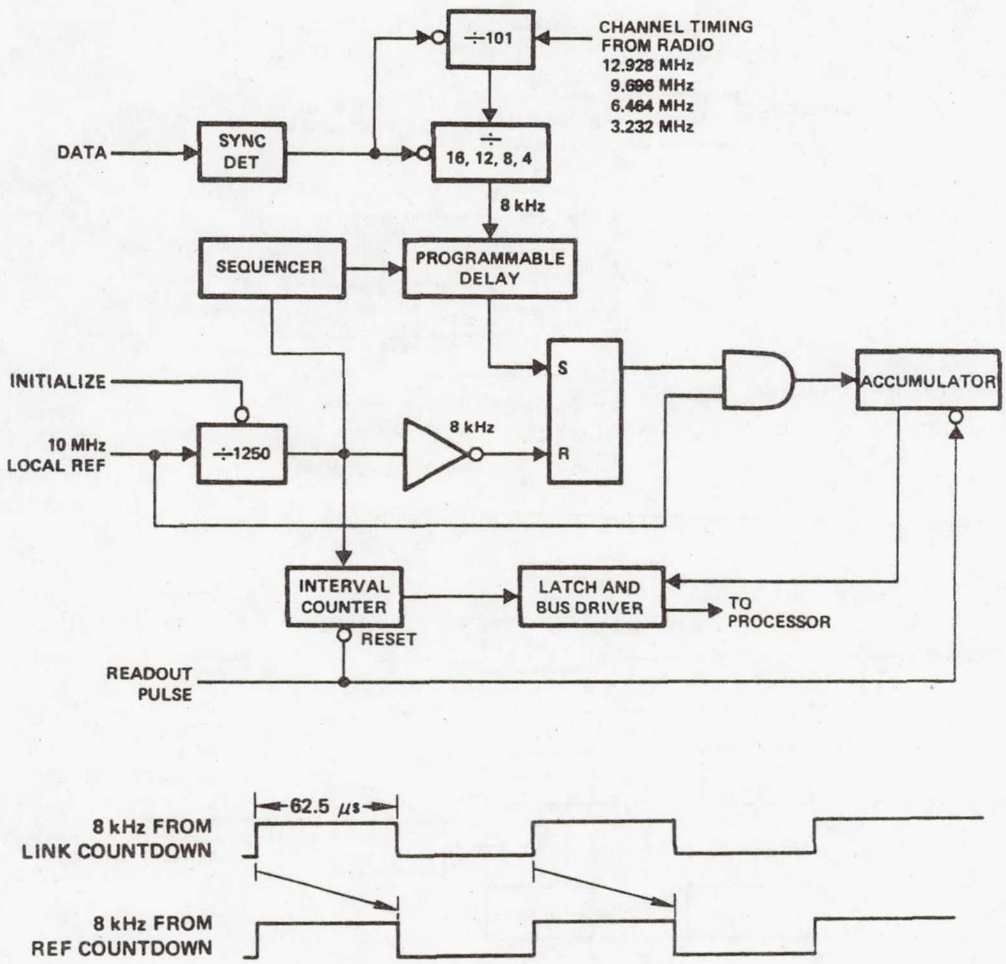


Figure 3. Phase Detector and Example Measurement

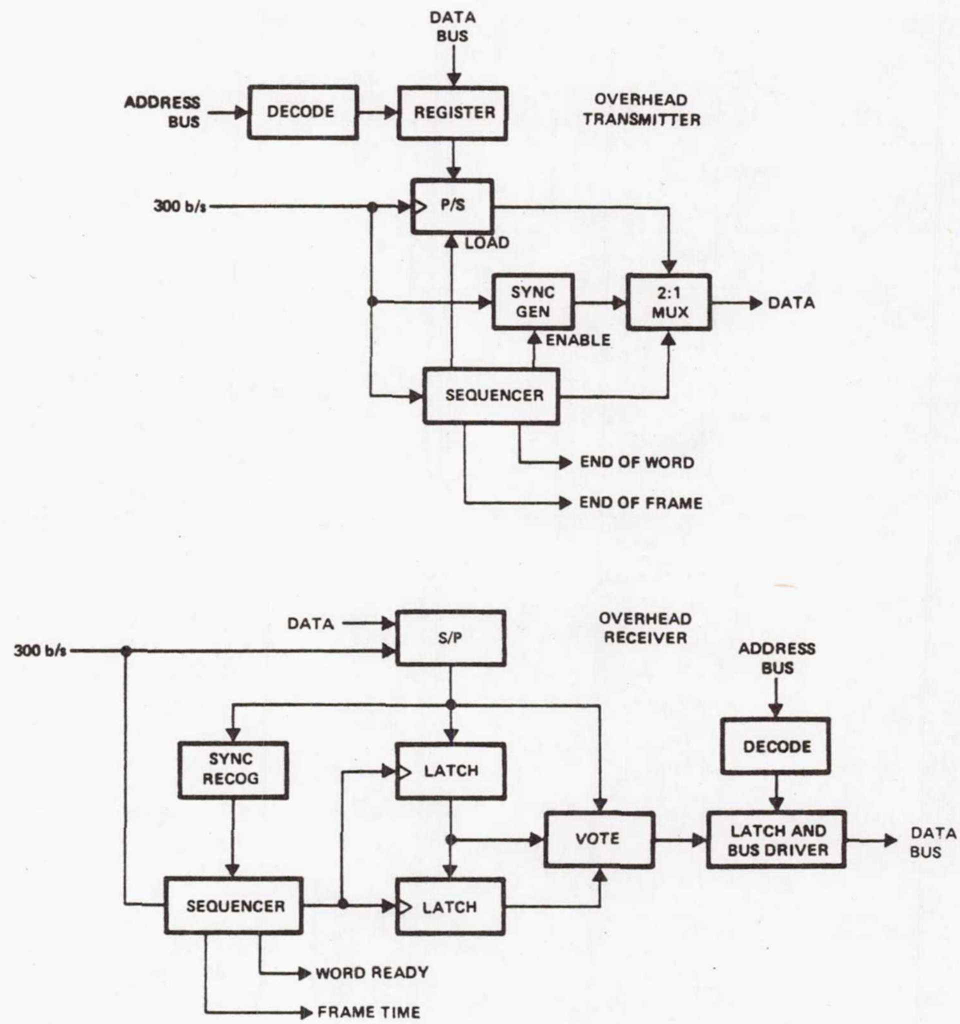


Figure 4. Overhead Interfaces

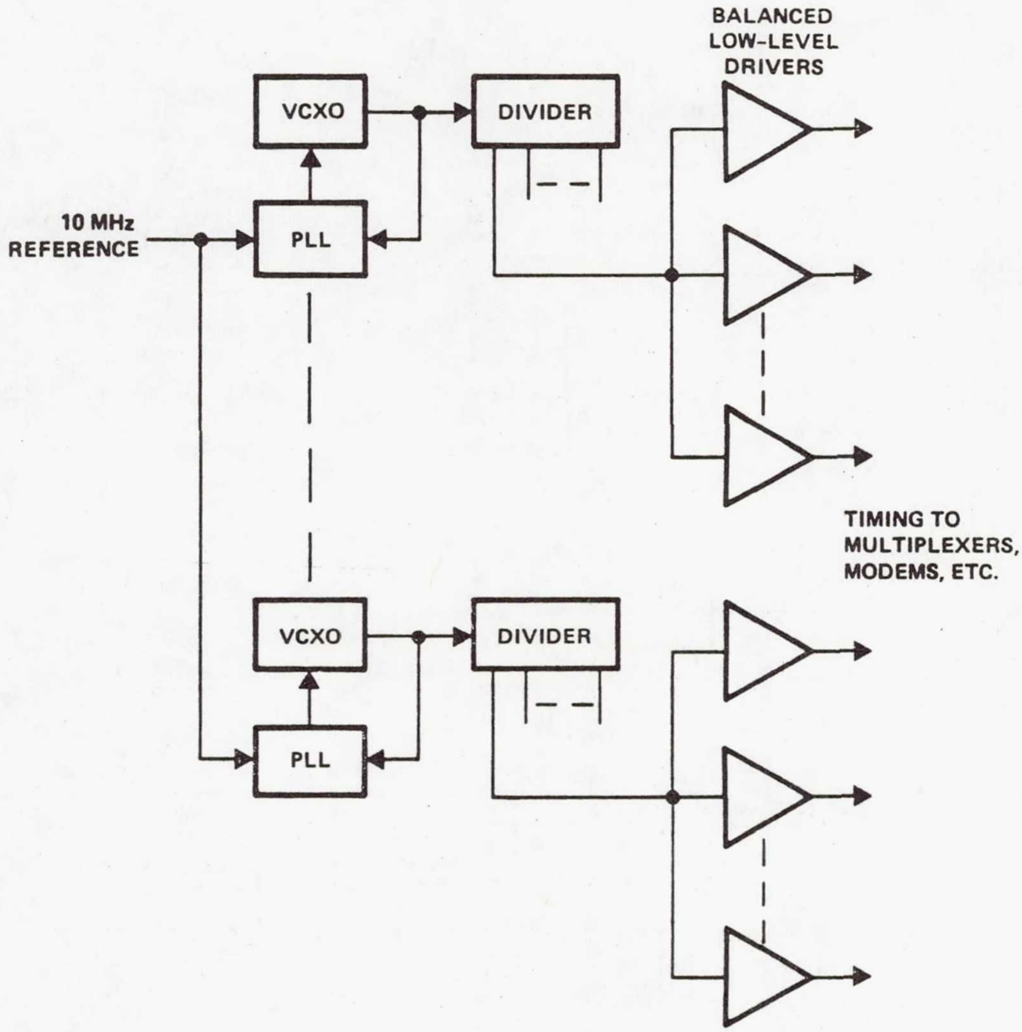


Figure 5. Frequency Synthesizers and Distribution Amplifiers

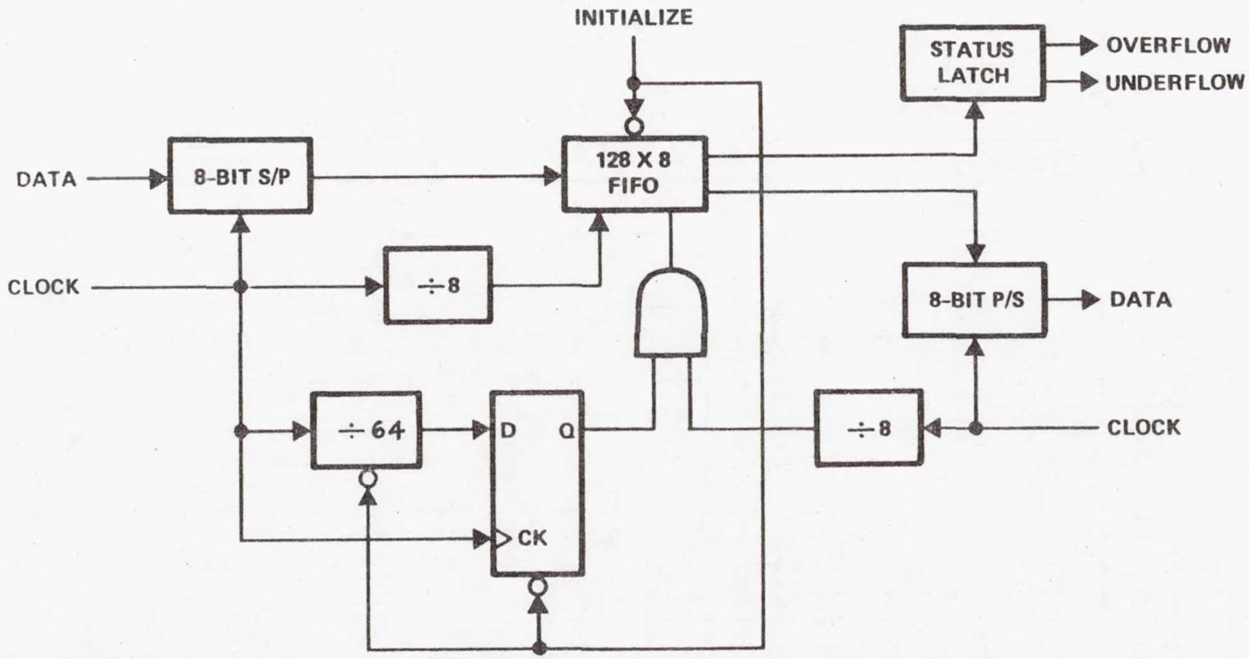
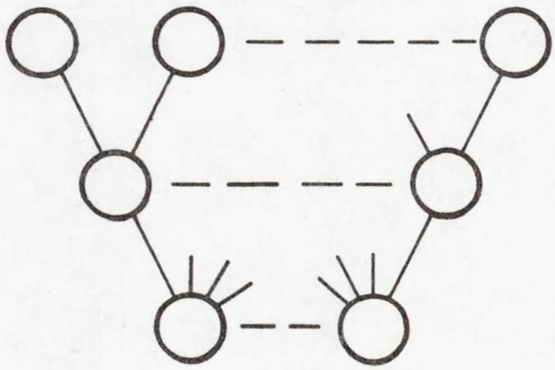


Figure 6. 1024-Bit Elastic Buffer



- SUPPORT YEARS – 20
- EQUIPMENT USAGE – 730.5 HOURS/MONTH (CONTINUOUSLY)
- PRICE 83B CHOOSES MTBF FOR DESIGNED ITEMS
- CATALOG VALUES USED FOR MTBF FOR PURCHASED ITEMS
- FOR RELIABLE ITEMS, NO SPARES AT INTERMEDIATE

Figure 7. LCC Assumptions

Table 1. Clock Frequencies and Maximum Number of Outputs Per Frequency

Clock Frequency Rates	Maximum Number of Outputs per Frequency Rate
16 kHz	25
32 kHz	10
56 kHz	15
64 kHz	10
128 kHz	15
256 kHz	15
512 kHz	10
1024 kHz	10
1544 kHz	60
2048 kHz	10
3.232 MHz	10
6.464 MHz	10
9.696 MHz	10
12.928 MHz	10

Table 2. Configurations for Costing

1-2.	Directed Control (Single-Ended)
3-7.	Mutual Control (Single-Ended)
8.	Directed Control, Double-Ended
9-10.	Mutual Control, Double-Ended
11.	Directed Control, Double-Ended, Independence of Measurement and Correction
12.	Directed Control, Double-Ended, Independence of Measurement and Correction, Phase Reference Combining
13.	Directed Control (Single-Ended), Self-Organizing
14.	Directed Control, Double-Ended, Self-Organizing
15.	Add Self-Organizing to Configuration 11
16.	Add Self-Organizing to Configuration 12
17.	Independent Clocks

Table 3. Software Requirements

Configuration		Lines	Bytes
1-2	Dir. Con., Sngl.	240	500
3-7	Mut. Con., Sngl.	360	760
8	Dir. Con., Dble.	440	920
9-10	Mut. Con., Dble.	630	1320
11	Dir. Con., Dble., IMC	480	1010
12	Dir. Con., Dble., IMC, PRC	1100	2310
13	Dir. Con., Sngl., SO	420	880
14	Dir. Con., Dble., SO	630	1320
15	Dir. Con., Dble., IMC, SO	670	1410
16	Dir. Con., Dble., IMC, PRC, SO	1300	2730
17	Ind. Clk.	0	0

Legend

Dir. Con. - Directed Control
Mut. Con. - Mutual Control
Ind. Clk. - Independent Clock
Dble. - Double-Ended
IMC - Indepencence of Measurement and Correction
PRC - Phase Reference Combining
SO - Self-Organizing
Sngl. - Single-Ended

Table 4. Basic Drawer Contents

Component	Ave./Box	Max/Box
Synthesizer 1	1	1
Synthesizer 2	1	1
Synthesizer 3	1	1
Synthesizer 4	1	1
Distribution Amp.	18	18
Dual Buffer	4	7
Motherboard	1.5	1.5
Supply (+5)	1	1
Supply (-5)	1	1
Supply (±12)	1	1
Drawer	1	1

Table 5. Disciplining Drawer Components

Component	Quantity
Processor	1
Controller	1
D/A	1
Supply (+5)	1 (Size Varies)
Supply (±12)	1
Supply (±15)	1
Motherboard	Varies
Drawer	1

Table 6. Disciplining Drawer Optional Components

	Sync. Ex.		Ph. Det.		OH Rcv.		OH Xmit		PROM
Configuration	Ave.	Max.	Ave.	Max.	Ave.	Max.	Ave.	Max.	
1-2	1	1	1	1	0	0	0	0	0
3-7	4	7	4	7	0	0	0	0	0
8	4	7	4	7	1	1	3	6	0
9-10	4	7	4	7	4	7	4	7	1
11	4	7	4	7	1	1	3	6	0
12	4	7	4	7	4	7	4	7	1
13	1	1	1	1	4	7	4	7	0
14	4	7	4	7	4	7	4	7	1
15	4	7	4	7	4	7	4	7	1
16	4	7	4	7	4	7	4	7	1
17	Not Applicable								

Legend

Sync. Ex. - Sync Extractor (1 card)
Ph. Det. - Phase Detector (2 cards)
OH Rcv. - Overhead Receiver (2 cards)
OH Xmit - Overhead Transmitter (1 card)
PROM - Additional 4K bytes memory (1 card)

Table 7. Software Costs (Thousands)

Configuration	Development	Maintenance
1-2 Dir. Con. (Sngl.)	3.6	24
3-7 Mut. Con. (Sngl.)	5.4	36
8 Dir. Con., Dble.	6.6	44
9-10 Mut. Con., Dble.	9.5	63
11 Dir. Con., Dble., IMC	7.2	48
12 Dir. Con., Dble., IMC, PRC	16.5	110
13 Dir. Con. (Sngl.), SO	6.3	42
14 Dir. Con., Dble., SO	9.5	63
15 Dir. Con., Dble., IMC, SO	10.1	67
16 Dir. Con., Dble., IMC, PRC. SO	19.5	130
17 Ind. Clk.	0	0

Table 8. Cost Breakdown for Configuration 16

	Qty	Dev	Prod	Support	Total	
Synth 1	1	29	72	185	287	
Synth 2	1	31	77	195	303	
Synth 3	1	46	71	188	305	
Synth 4	1	31	75	193	298	
Dist Amp	18	14	399	1403	1816	
Elas Buf	4	21	162	519	702	
Backplane	1	65	458	2	525	
Power 3	1	0	116	138	255	
Power 4	1	0	137	144	281	
Power 8	1	0	24	75	99	
Drawer	1	2	647	163	811	
I&T*	1	6	430	96	532	
Subtotals		245	2668	3301	4214	(Basic Box)
PC1	4	18	103	305	426	
PC2	4	19	103	329	450	
Syncex	4	33	117	406	556	
OHRCV1	4	28	141	454	623	
OHRCV2	4	20	93	288	401	
OHXMTR	4	23	105	360	488	
PROM	1	14	68	132	214	
8080	1	11	60	145	216	
D/A	1	48	90	224	361	
Timer	1	24	59	142	225	
Power 5	1	0	196	154	351	
Power 9	1	0	85	131	216	
Power 10	1	0	37	115	152	
Mother	1	49	375	2	427	
Drawer	1	0	647	163	810	
I&T*	1	6	157	324	487	
Subtotals		293	2436	3674	6403	(Disc Box)
STAL	1	3	1281	129	1414	
I&T*	1	12	416	142	570	
Totals		553	6801	7246	14601	

*Note: I&T means Integration and Test

Table 9. Life-Cycle Costs (In Thousands)

Configuration	Hardware	Software	Total
1-2	11346	28	11374
3-7	12454	41	12495
8	13318	51	13369
9-10	14601	73	14674
11	13318	55	13373
12	14601	127	14728
13	13361	48	13409
14	14601	73	14674
15	14601	77	14678
16 (Crystal)	14601	150	14751
16 (Cesium)	21635	150	21785
16 (Rubdium)	19229	150	19379
17 (Cesium)	13704	0	13704
17 (Rubdium)	11306	0	11306

Page intentionally left blank

CLOCK PERFORMANCE AS A CRITICAL PARAMETER IN NAVIGATION SATELLITE SYSTEMS

Richard J. Anderle
Naval Surface Weapons Center, Dahlgren, Virginia 22448

ABSTRACT

The high performance of available oscillators has permitted the development of invaluable navigation and geodetic satellite systems. However, still higher performance oscillators would further improve the accuracy or flexibility of the systems.

INTRODUCTION

Oscillator performance is a critical factor in the operation of the Navy Navigation Satellite System (NAVSAT) and of the NAVSTAR Global Positioning System (GPS). It is also an important element in the geodetic applications of these systems. The NAVSAT system is based on Doppler observations of satellites at 1000 km altitude. While the GPS system is based on simultaneous range observations to four satellites at altitudes of 20,000 km, it is useful to think of the computation of the ephemerides of the satellites as being based upon Doppler data also. The reason for choosing this interpretation stems from the fact that it is desirable to base the computation of the ephemeris on several days of observations in order to minimize the uncertainties in the computed orbit period and solar radiation parameters. Over a five day period, an error of one part in 10^{-13} in oscillator frequency would produce an error in time of 43 ns, or 12 m in range, since the range is based on the measured travel time of signals propagating at the speed of light. As will be shown below, analysis of Doppler data during the five day period would give range to the satellite which is accurate to better than a meter using the same oscillator. It is possible to use the range information directly while still accounting for the oscillator instability either in a sequential processor by introducing process noise or in a batch processor by using a correlated weight matrix. While these alternatives are mathematically more rigorous than the conversion of range data to Doppler data, the techniques fundamentally weaken the accuracy of relative range measurements made at widely spaced times, tending to approach the Doppler interpretation of the data.

TIME TAGGING

Both the NAVSAT and GPS systems require time tagging of observations and of ephemeris data to sufficient accuracy to allow interpolation in the relative positions of the satellite and observer to the desired accuracy. Since the relative accuracy of the satellite and observer is about 5 km/sec, an uncertainty in the time tag of 0.2 ms would produce a relative position error of one meter. In both navigation systems, one or more ground oscillators is adopted as a standard, the satellite oscillator is calibrated against the standard, and other ground clocks are calibrated against the satellite clock. Therefore, the satellite oscillator must be sufficiently stable to maintain the desired accuracy of the clock epochs over the time period of several days used for the clock rate determination and prediction. A 0.2 ms accuracy objective over a five day period requires an oscillator stability of 5 parts in 10^{10}.

RANGE MEASUREMENTS

The most stringent requirement on oscillator performance arises from ground measurements of the time of arrival of signals generated from oscillators in the GPS satellites. The GPS system is based on ranges computed by multiplying the travel time of the signals by the velocity of light. The effect of oscillator instability on the computed ranges was referred to in the first paragraph in connection with the determination of the orbits of the GPS satellites. Inverting the calculation, if satellite and ground timing systems were to be maintained to an accuracy corresponding to a one meter range accuracy over a five day time period, oscillator stabilities of eight parts in 10^{15} would be required. The GPS system is able to meet navigation requirements with satellite oscillators which are an order of magnitude poorer because of looser tolerances on range accuracy and shorter fit and prediction intervals for the time signals. The epoch errors of the ground clock are determined each time a navigation fix is obtained by measuring the apparent travel time of signals from four satellites and solving for the clock correction and the three components of the observer's position. Therefore the only requirement on the oscillator in the receiver is to permit interpolation of signals from the satellite to the same epoch for those receivers which do not make simultaneous observations to the four satellites (Hill, 1978). The range computed from the travel time prior to correction of the observer's clock is referred to as a "pseudo-range."

GEOMETRIC DILUTION OF PRECISION

In considering the requirements for oscillator stability, the measurement errors produced by clock uncertainties must be transformed to errors in the position of the observer. Positions based on the pseudoranges to four satellites are about a factor of three worse than the

measurement errors for the typical geometric configuration of GPS satellites. The ratio of the position error to the measurement error is referred to as the "Geometric Dilution of Precision (GDOP)" (more precisely in this context, "Position Dilution of Precision (PDOP)" (Milliken and Zoller, 1978)). It is simply the average standard error in position corresponding to unit weight for the observations. The GDOP and the effects of oscillator instability on Doppler positioning cannot be summarized as concisely. Before discussing these topics, the conventional interpretation of Doppler data and common terminology will be reviewed.

SENSITIVITY OF DOPPLER DATA TO CLOCK PERFORMANCE

The observed frequency is given to first order by:

$$f = f_s - \frac{f_s}{c}\dot{r}$$

where f_s is the frequency emitted by the satellite, c is the velocity of light, and $\dot{r}$ is the relative velocity of the satellite with respect to the observer. The received frequency is normally mixed with a reference frequency, f_R, in the receiver:

$$f_B = f_R - f_s + \frac{f_s}{c}\dot{r}$$

and the beat cycles are counted over specified time intervals:

$$N_c = \int_{t_1}^{t_2} f_B \, dt = \int_{t_1}^{t_2} (f_R - f_s + \frac{f_s}{c}\dot{r}) \, dt$$

so that

$$N_c = (f_R - f_s)(t_2 - t_1) + \frac{f_s}{c}(r_2 - r_1)$$

Some receivers measure $(t_2 - t_1)$ for fixed N_c, some count N_c for fixed $(t_2 - t_1)$, and some count integer N_c in fixed $(t_2 - t_1)$ and read out the clock at the time corresponding to the integer N_c. Many receivers continue counting as the measurements are made and recorded, so that the measurements at the i^{th} data point can be written as $r_i - r_o$ rather than as $r_i - r_{i-1}$. In such cases, the Doppler measurements during a satellite pass can be represented as ranges subject to an unknown range base, r_o, rather than as uncorrelated range differences. Biased range representation yields a better GDOP than uncorrelated range differences, as will be shown later. The offset frequency $f_R - f_s$ varies among receivers. For the Navy Navigation Satellites the offset ratio $(f_R - f_s)/f$ is 80×10^{-6}; for the NAVSTAR Geodetic Receiver (Anderle, 1978c) the offset $(f_R - f_s)$ is 28.75 KHz. To first order, oscillators make two contributions to the range error:

$$\delta_1(r_2 - r_1) = \frac{c}{f_s}(f_R - f_s)\delta t$$

$$\delta_2(r_2 - r_1) = \frac{c}{f_s}(t_2 - t_1)\delta(f_R - f_s)$$

For the above frequency offsets, the first equation establishes the time interval accuracy required per meter precision in range difference as 40 μs for NAVSAT and 200 μs for GPS. The second contribution to the range difference error imposes more severe requirements on the oscillator. The time interval between the first and last time in a satellite pass is about 1000 s for NAVSAT and 30,000 s for GPS. Therefore, the fractional frequency stability required per meter precision over these intervals is $3x10^{-12}$ for NAVSAT and $1.1x10^{-13}$ for GPS.

INFORMATION CONTENT OF A DOPPLER PASS

Direct conversion of the Doppler errors discussed in the previous paragraph to errors in station position is not useful because Doppler data for a single satellite pass does not provide enough information to permit accurate determination of all three components of station position. Therefore, GDOP is usually calculated for the two position components which are well determined. The effects of errors in these two components on the calculated frequency are illustrated in figure 1. On a non-rotating earth, the Doppler frequency (which is proportional to the range difference) is zero when the satellite reaches its point of closest approach to the observer and has the shape shown by the upper curves in the figure. The offset between the satellite and station frequency standards is easily determined since the Doppler frequency, or calculated range difference per unit time, is equal and opposite in sign at the times of rise and set of the satellite above the station horizon. If the satellite position is known, then an error in the observer's position parallel to the satellite velocity vector at closest approach will produce calculated range differences which are displaced in time as shown by the broken curve in the upper left figure, and bell shaped residuals as shown in the lower left figure. This component of station position determined from a pass of Doppler data is referred to as the "tangential" or "along track" component of position. If the assumed station position is closer to, or further from, the satellite at the time of closest approach, the Doppler curve, or range differences, will have a steeper or shallower slope as shown in the upper right hand part of figure 1. The residuals will be anti-symmetric as shown in the lower right hand part of the figure, and define the location of the station alonw the range vector to the satellite at the time of closest approach (the "range" component of station position). A tropospheric refraction bias will also produce anti-symmetric residuals, but the effect will be greatest at the times of rise and set of the satellite

and decrease rapidly at the higher elevation angles. The third component of station position is not defined for an emitter on a linear path and a non-rotating earth, since rotation of the receiver about the emitter path at a fixed distance from the emitter will not change the Doppler curve. While the solution for three components of station position is not singular for the curved satellite path and a rotating earth, the standard error for the third component of station position is orders of magnitude larger than those for the tangential and range component of station position in the plane defined by the range vector to the satellite and the relative velocity vector of the satellite at the time of closest approach, providing no useful information for navigation or geodetic applications. Therefore in order to determine three components of station position, the satellite should be observed on a pass to the left and a pass to the right of the station so that the range components can be used to triangulate station height and the horizontal component normal to the satellite track (longitude for the polar Navy Navigation Satellites). In order to determine a navigator's latitude and longitude from a single pass of Doppler data, the height of the observer must be known; nevertheless the longitude is ill-determined for polar satellite passes crossing the station's zenith. Since the angular velocity of GPS satellites is only twice the rate of earth's rotation while the angular velocity of the NAVSAT satellites is ten times the rate of earth's rotation, it is not clear whether the information content of a GPS Doppler pass is so ideally contained in the range/tangential position components of station position as it is for NAVSAT data. Nevertheless, the same interpretation has been applied to GPS data as a result of the availability of the computer programs and the lack of a better diagnostic tool. Actual orbit determinations and geodetic station position calculations are based on a least squares fit of the parameters of the solution to the aggregate of the Doppler data, not to the position components calculated for diagnostic purposes.

EFFECT OF CLOCK PERFORMANCE ON POSITIONS DETERMINED FROM NAVSAT DOPPLER DATA

It was mentioned earlier that Doppler observations from most receivers can be treated as either range difference data or as range data subject to an unknown bias. Figures 2 and 3 show the uncertainty in the determination of the tangential and range components of the position of the observer, respectively, corresponding to a 10 cm random error in range or range difference data. Figure 2 indicates that the GDOP for the tangential component of position varies from one to four for biased range data and from three to ten for range difference data for elevation angles to the satellite at closest approach from 90 to 20 degrees. Figure 3 reveals that the GDOP for the range component of position varies from one half to two for biased range data and from three to seven for range difference data for these elevation angles. The figures are based on the assumption that the tropospheric refraction is known perfectly and the offest in frequency between the oscillators in the

satellite and the receiver is completely unknown but stable during the pass. Uncertainties in tropospheric refraction must actually be considered in precise computations. Introduction of a scale bias for refraction does not affect the standard error in tangential position. The effect on the range component of position depends on the relative magnitudes of the random error of the Doppler observations and the uncertainty in the a-priori refraction data; for typical values of the quantities, the standard error in range component based on range difference data is not significantly affected while that for biased range data is increased markedly percentage-wise, although it always remains smaller in magnitude than that for range difference data. Since the random error of measurement for the better receivers is less than 5 cm, the precision of the Doppler receivers is quite good. However, the effects of the instability of oscillators used in the receivers produces larger errors in position. Specifications of the stability of two oscillators used in NAVSAT Doppler receivers are 1×10^{-11} and 6×10^{-12} for averaging times of interest (30 to 1000 seconds). Simulations of position accuracies attainable with these oscillators and an oscillator with a stability of 2×10^{-13} were conducted by Monte Carlo methods. Doppler observations corresponding to frequency variations expected for each of these oscillators and a random error of 3 cm were synthesized for six passes for each of five pass geometries, and the components of station position were computed for each pass. The rms of the six sample errors for the tangential and range components is plotted in figures 4 and 5, respectively, versus the elevation angle to the satellite at closest approach. Note that the position component errors are about 30 times larger than those due to random error for the specifications of the oscillators used with this equipment regardless of whether the data is represented as biased range data or as range differences. The oscillator stability of 2×10^{-13} which has been achieved for rubidium oscillators over these averaging times, yields position errors reasonably close to those expected from the random error of measurement. Irregularities in the curves are probably due to sampling errors in this limited Monte Carlo simulation. The rubidium oscillator is inconveniently large in size for use with the portable Doppler receivers in some applications.

EFFECT OF CLOCK PERFORMANCE ON POSITIONS DETERMINED FROM GPS DOPPLER DATA

Results of computations of GDOP for Doppler observations of the GPS satellites for biased range and range difference data are given in figure 6 for the range component of position. The curves for the tangential component of position are similar. Results for various data sampling strategies are given for the range difference representation of data while the curves for biased range data are proportional to the square root of the sampling interval. Note that the GDOP varies from about one to ten for the different cases for pass lengths greater than 15,000 seconds. Shorter pass lengths probably need not be considered

due to the spacing of the satellites in the GPS constellation. Since most GPS receivers are designed to achieve 1 cm precision in Doppler data, the curves imply a high precision in position. However, a very high oscillator stability would be required to achieve these precisions. Simulations similar to those conducted for NAVSAT conditions were also conducted for GPS conditions to determine the effect of oscillator stability on the accuracy of station positions. Data were simulated for the oscillator stability corresponding to the curve labeled "Test A" on figure 7. This curve is close to that for a cesium oscillator, just a little poorer than that measured by the Naval Observatory for the cesium oscillator used in the NAVSTAR Geodetic Receiver. The rms of each position component error obtained from the simulated data is given in figure 8. Only pass lengths longer than 15,000 seconds were considered. These errors are five to fifty times worse than those expected from the random error of observation. Attempts to account for frequency variations by introducing a frequency drift parameter produced still larger errors in computed station position. However, this figure illustrates the point made in the first paragraph that the Doppler technique can be used to determine the range to the satellite to better than a meter accuracy for satellite passes separated by any time interval.

RELATIVE STATION POSITIONING

Even considering the effects of oscillator instability, the errors in computed station positions discussed in the previous sections are smaller than the errors in computed satellite positions except for low elevation angle passes. However, the higher receiver accuracy is desirable for geodetic applications since the accuracy of the computation of the relative position of stations observing the satellite simultaneously is not significantly affected by errors in the satellite position if the distance between the stations is small compared to the height of the satellite (Anderle, 1978a). Similarly, errors due to the satellite oscillator can be expected to be cancelled under these circumstances. The potential for the determination of the relative positions of stations to centimeter accuracy has attracted the attention of geophysicists studying crustal motion. Since the determination of relative station position also negates the requirement for accurate times of emission of the ranging signals from the GPS satellites, near-simultaneous pseudo-range measurements from two stations to four satellites can be used to make an interferometric solution for the relative position of the stations (Anderle, 1978b, MacDoran, 1978). However, a high gain antenna or a high redundancy of observations is required to reduce the random range error which is about a meter for a wide beam antenna. In this application, oscillator requirements are modest since accurate time intervals are only required to interpolate non-synchronous but high data rate data.

SUMMARY

The high performance of available oscillators has permitted the development of invaluable navigation and geodetic satellite systems. However, still higher performance oscillators would improve the accuracy of flexibility of the systems. Oscillator requirements per meter position error are listed in figure 9 for the various aspects of navigation systems discussed in this report. A GPS oscillator stability of 10^{-15} over five days would simplify the orbit determination and prediction functions. Highly portable low cost oscillators with a stability of 10^{-14} for averaging times of eight hours would permit monitoring of crustal motion daily with GPS Doppler receivers. Oscillators the same size and reasonably close to the cost of current quartz oscillators but with a stability closer to 10^{-13} at an averaging time of 1000 seconds would allow more rapid determination of relative station postions from NAVSAT data and more accurate orbit determination. Clearly clock performance is a critical parameter in navigation satellite systems.

ACKNOWLEDGEMENTS

Simulations of the effects of oscillator stability on the position accuracy obtainable from Doppler observations of the NAVSAT and GPS satellite systems were conducted by Ronald Smith of the Naval Surface Weapons Center.

REFERENCES

1. Anderle, Richard J., "Geodetic Applications of the NAVSTAR Global Positioning System," preprint of paper presented at the Second International Symposium on Problems Related to Redefinition of North American Geodetic Networks, Washington, D.C., April 1978a.

2. Anderle, R. J., "Application of Global Positioning System to Dedetmination of Tectonic Plate Movements and Crustal Deformations," preprint of paper presented at the Ninth Geodesy/Solid-Earth and Ocean Physics (GEOP) Research Conference, Columbus, Ohio, October 1978b.

3. Anderle, R. J., "Accuracy of Geodetic Solutions based on Doppler Measurements of the NAVSTAR Global Positioning System Satellites," Bulletin Geodesique, in press 1978c.

4. Hill, R. W., "Applications of the NAVSTAR Global Positioning System to Marine Geodesy," preprint of paper presented at the International Symposium on Interaction of Marine Geodesy and Ocean Dynamics, Miami, Florida, October, 1978.

5. MacDoran, P. F., "Mobile Long Baseline Interferometry," presented at the Ninth Geodesy/Solid-Earth and Ocean Physics (GEOP) Research Conference, Columbus, Ohio, October 1978.

6. Milliken, R. J. and C. J. Zoller, "Principle of Operation of NAVSTAR and System Characteristics," Navigation, 25(2), Summer 1978.

INFORMATION CONTENT OF DOPPLER PASS

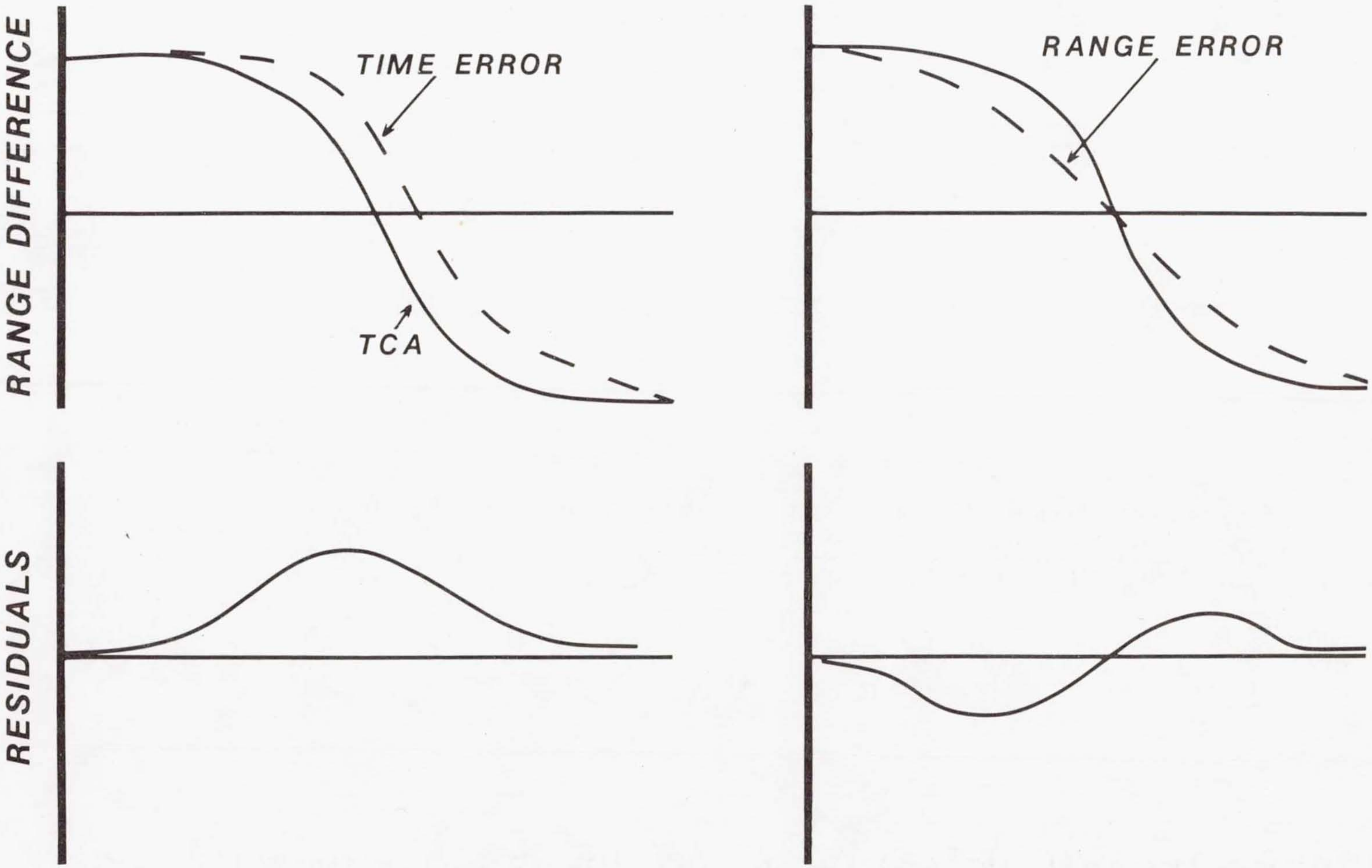


FIGURE 1

STANDARD ERROR IN TANGENTIAL COMPONENT OF POSITION FOR 10 CM RANDOM ERROR OF OBSERVATION NAVSAT PASSES

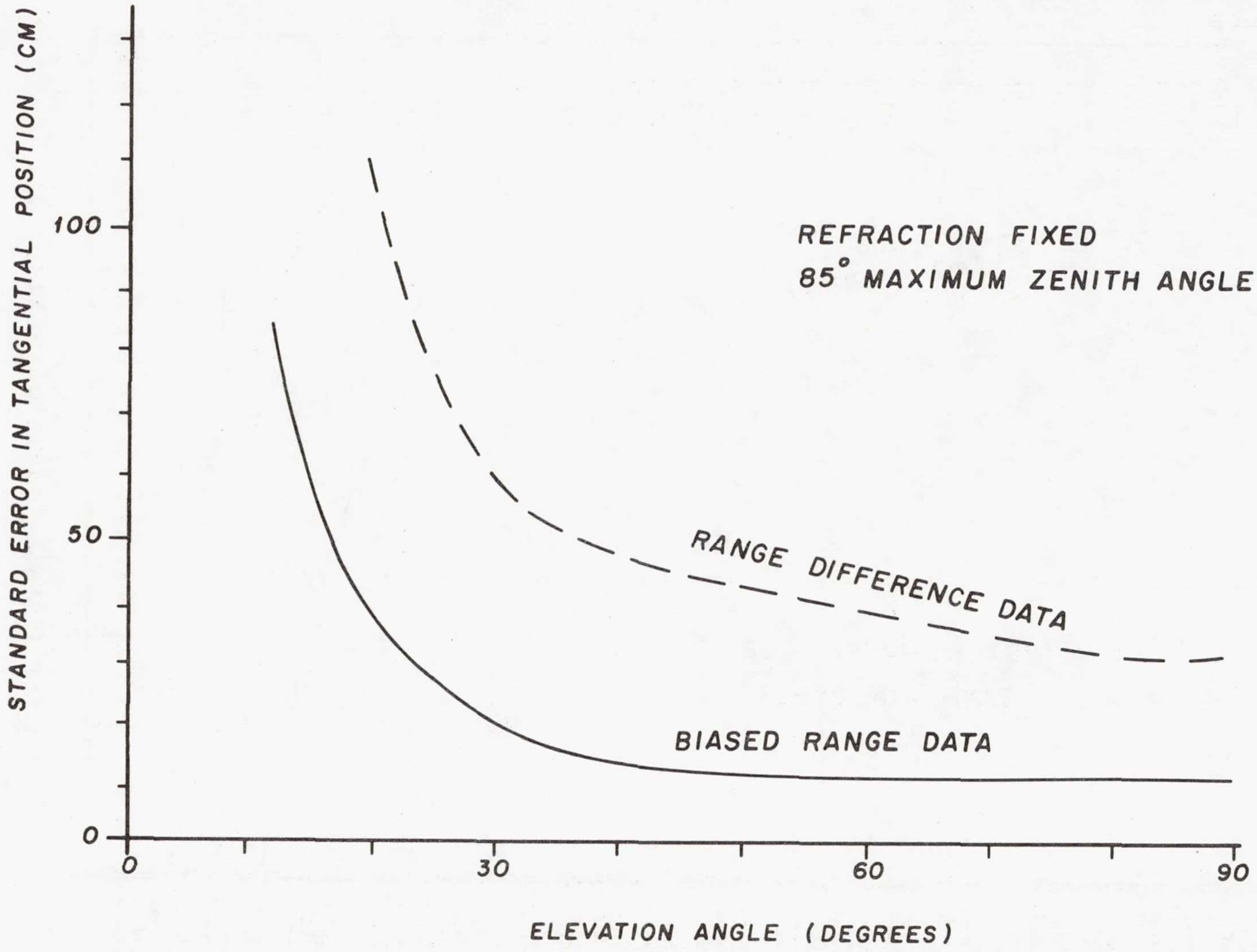


FIGURE 2

STANDARD ERROR IN RANGE COMPONENT OF POSITION
FOR 10 CM RANDOM ERROR OF OBSERVATION NAVSAT PASSES

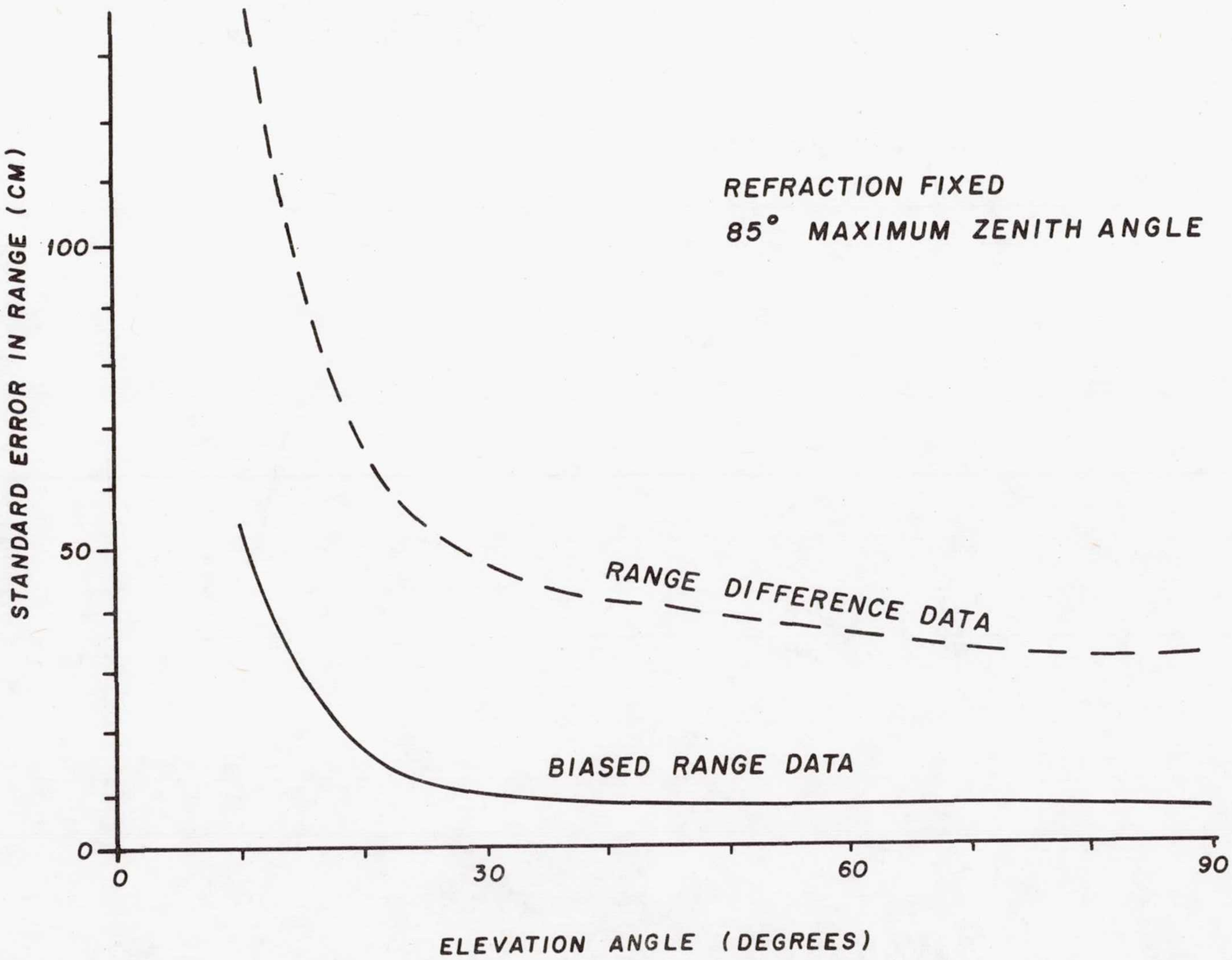


FIGURE 3

EFFECT OF OSCILLATOR ERROR ON TANGENTIAL POSITION COMPONENT

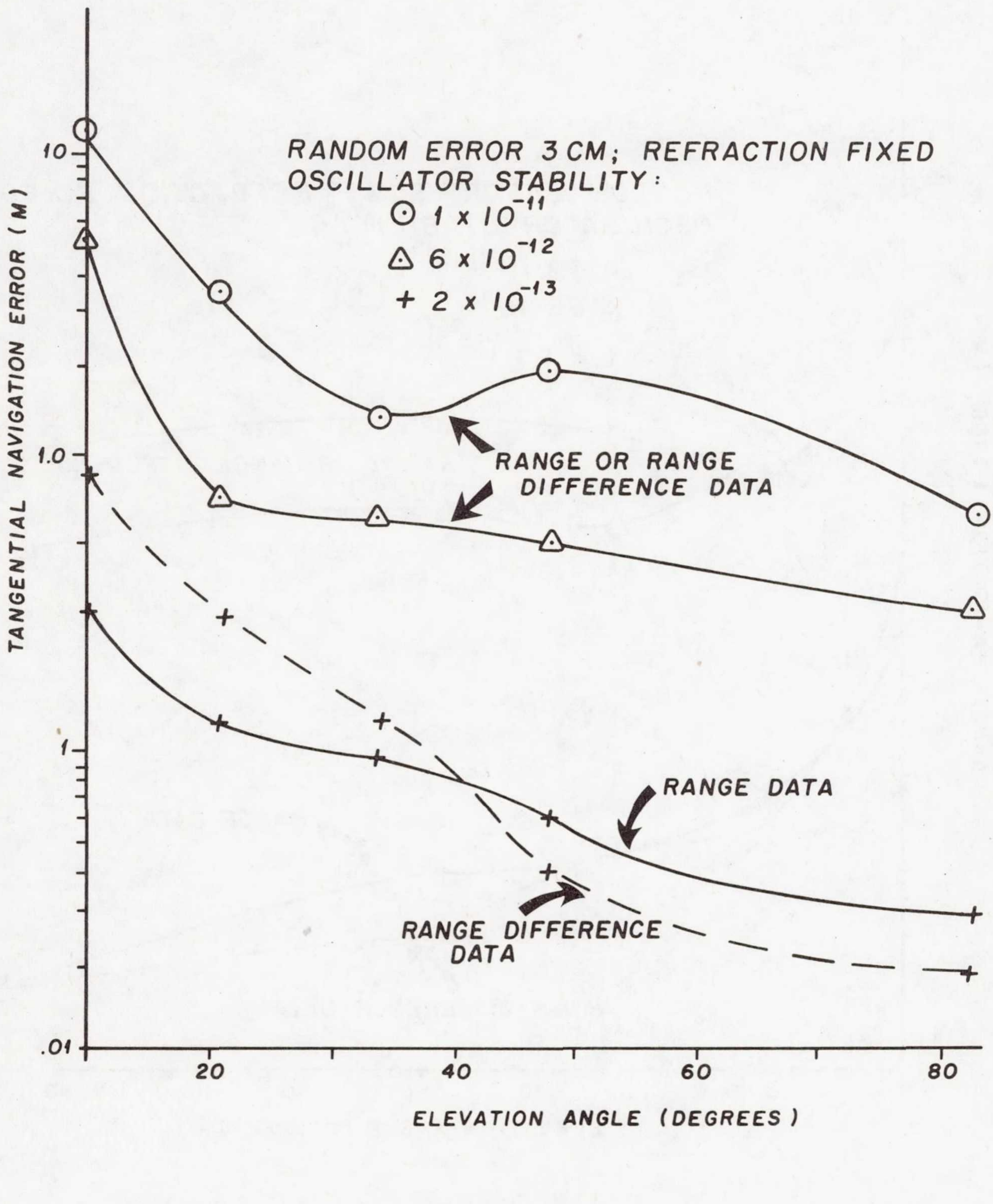


FIGURE 4

EFFECT OF OSCILLATOR ERROR
ON RANGE POSITION COMPONENT

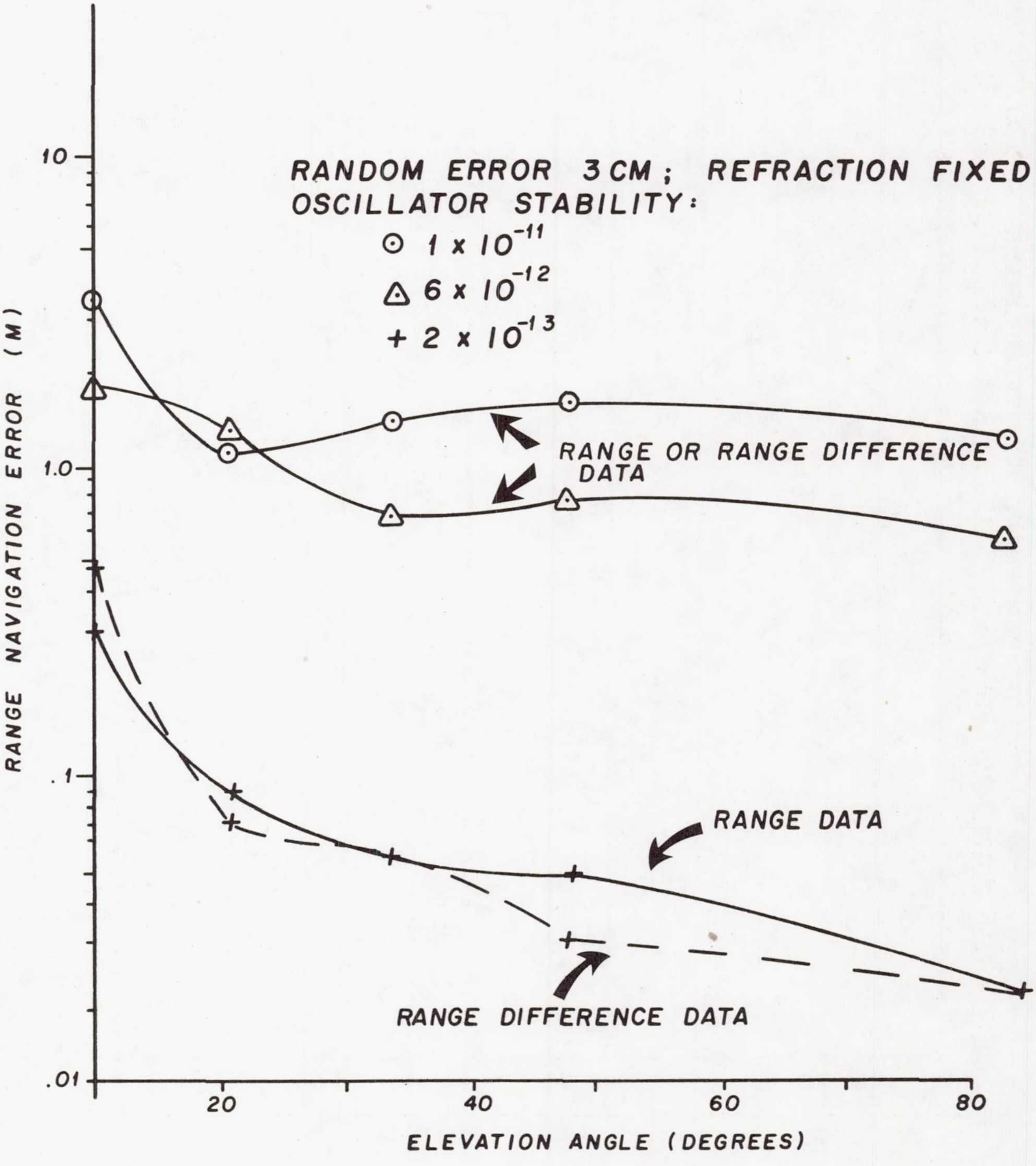


FIGURE 5

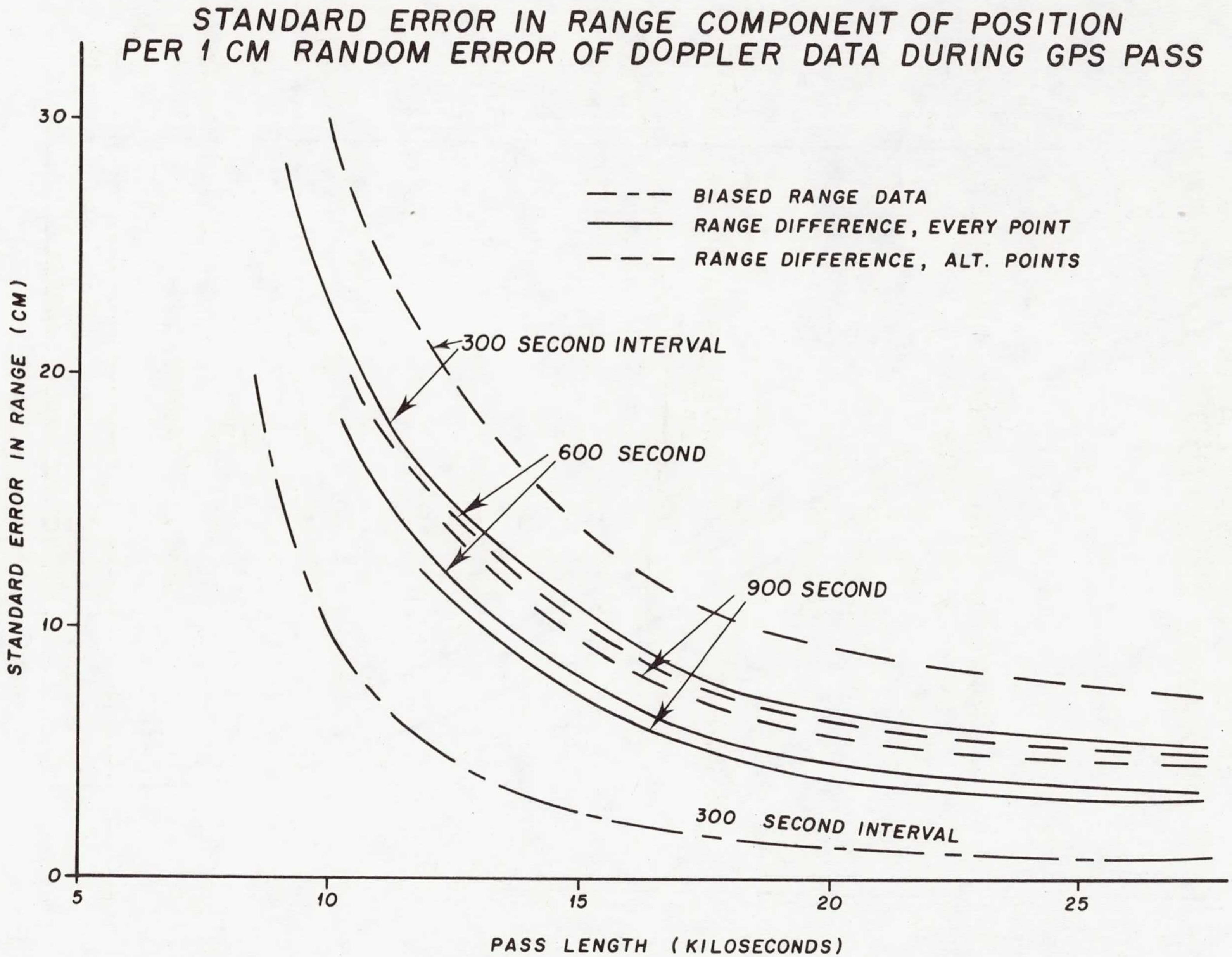
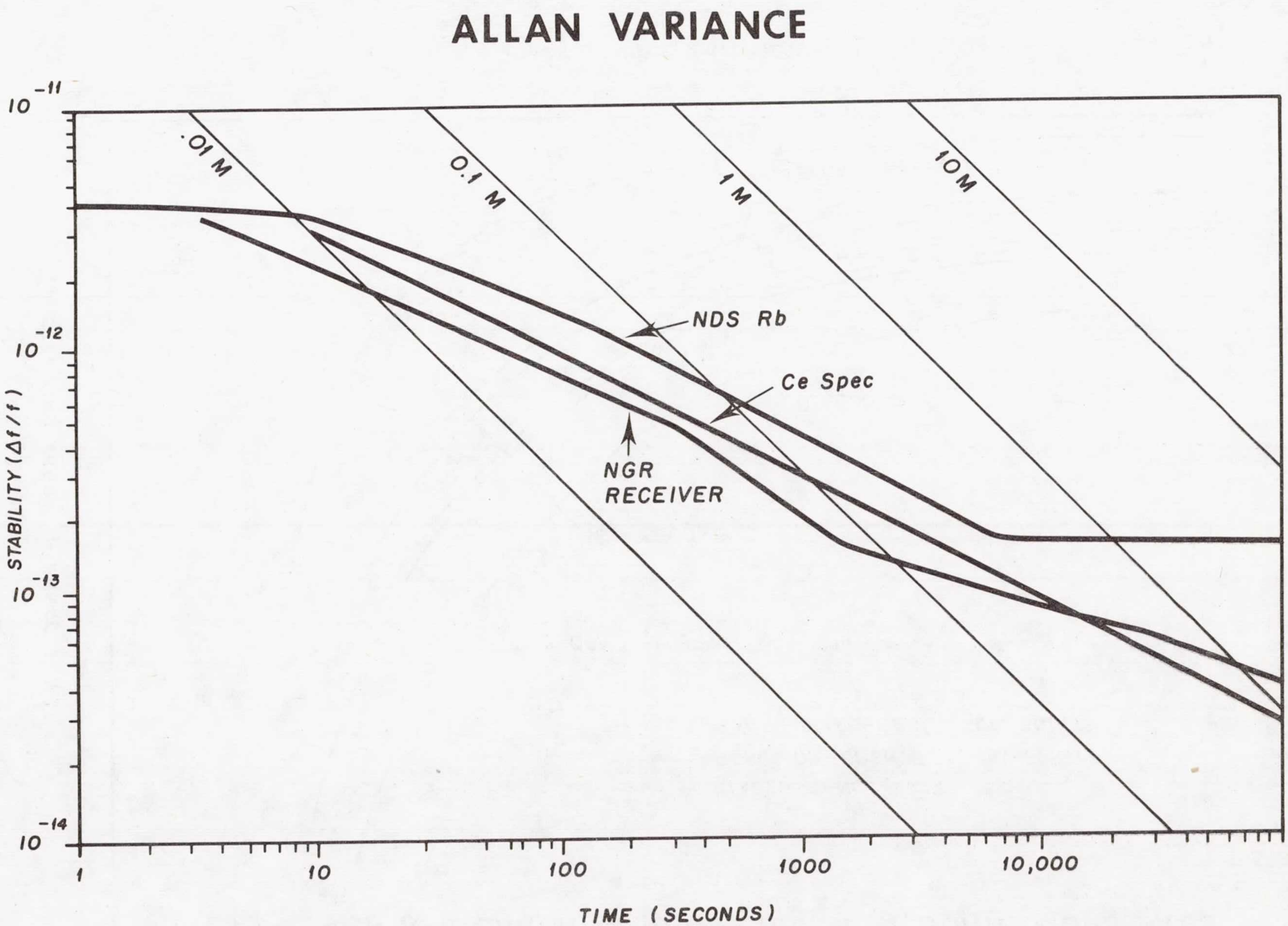

STANDARD ERROR IN RANGE COMPONENT OF POSITION
PER 1 CM RANDOM ERROR OF DOPPLER DATA DURING GPS PASS
BIASED RANGE DATA
RANGE DIFFERENCE, EVERY POINT
RANGE DIFFERENCE, ALT. POINTS
300 SECOND INTERVAL
600 SECOND
900 SECOND
300 SECOND INTERVAL
STANDARD ERROR IN RANGE (CM)
PASS LENGTH (KILOSECONDS)
0
10
20
30
5
10
15
20
25

FIGURE 6



ALLAN VARIANCE
10^{-11}
10^{-12}
10^{-13}
10^{-14}
STABILITY ($\Delta f/f$)
.01 M
0.1 M
1 M
10 M
NDS Rb
Ce Spec
NGR
RECEIVER
1
10
100
1000
10,000
TIME (SECONDS)

FIGURE 7

506

RMS POSITION COMPONENT ERRORS (RADIAL/TANGENTIAL) IN CM
DUE TO OSCILLATOR INSTABILITY
(PASSES LONGER THAN 15,000 SECONDS)

DATA TYPE	DATA INTERVAL (SECONDS) 300	600	900
RANGE	45/42		44/35
RANGE DIFFERENCE (CONTINUOUS)	36/38	30/32	29/43
RANGE DIFFERENCE (ALTERNATE POINTS)	41/56	48/62	49/73

FIGURE 8

ROLE OF CLOCK PERFORMANCE
IN NAVIGATION SATELLITE SYSTEMS

TIME TAGGING	POSITION ERROR*	CRITICAL T	OSCILLATOR REQUIREMENT PER M POSITION ERROR	
			EPOCH	FREQUENCY
EPHEMERIS DETERMINATION	$V_s \delta\tau,\ V_s \int_0^T \frac{\delta f}{f} dt$	6 hours	100 μs	10^{-9}
SATELLITE - NAVIGATOR		24 hours	100 μs	2.5×10^{-9}
RANGE POSITIONING	$3c\delta\tau,\ 3c \int_0^T \frac{\delta f}{f} dt$	24 hours	1 ns	1.3×10^{-9}
DOPPLER POSITIONING				
NAVSAT	$2V_3\, \delta\tau\ ,\ 2c \int_0^T \frac{\delta f}{f} dt$	.3 hours	50 μs	1.5×10^{-12}
GPS		8 hours	50 μs	5×10^{-14}

*The coefficients "2" or "3" represent typical ratios of position error to measurement error.

FIGURE 9

QUESTIONS AND ANSWERS

DR. VICTOR REINHARDT, NASA Goddard Space Flight Center:

Can you define the term, g-dop?

DR. ANDERLE:

Well, loosely speaking, it is the position accuracy per unit measurement accuracy. The detailed definitions of the GPS are given in the last issue of "Navigation". A number of conditions are involved: horizontal, g-dop, position-dop, a number of those terms, but fundamentally, it is, loosely speaking, position accuracy per unit measurement accuracy.

DR. IVAN NURUR, Ohio State:

In these biased ranges, did you assume that these ranges are independent from each other on a given pass, or did you consider correlations between them?

DR. ANDERLE:

Each measurement I assume is essentially independent. The only common bias is the range bias for the pass, but each biased range is independent of the preceding one.

MR. MIKE MCCONAHY, Johns Hopkins University, Applied Physics Lab:

Would you like to comment, Dick, on what you think the potential of GPS is for geophysical studies, in view of what you now know?

DR. ANDERLE:

I have addressed that in a number of papers and there are a number of ways, I think, of achieving centimeter accuracies in fairly short time spans with better oscillators. This doppler receiver would do it, and with the existing receivers, depending upon how biases work out and depending on averaging times, it is theoretically possible to get a centimeter that way also. There are a number of other proposals that have been made for using GPS in a VLBI mode as another technique. So, there are four or five different approaches to using GPS for geophysical applications. There is a question of what the equipment cost; you know, which one would have the least cost, the fastest operation, how the various system classes would work out in each respective application. I don't have any doubt that one of them will work for centimeter accuracy at some acceptable cost.

Speaker (unheard)

DR. ANDERLE:

I am sorry. When I talk about those accuracies, I am talking about relative positioning; I am not talking about absolute positioning. But, there are two stations equipped with these things, in getting relative positions.

DR. WILLIAM MURPHY, Rockwell:

You might have made this point clear but it wasn't clear to me. When you were speaking about clock performance, were you talking about stability or accuracy?

DR. ANDERLE:

In terms of absolute time tags. I never talked in terms of absolute time tags, absolute epochs because, as I say, we adopt some ground station as a standard and time tags are with respect to that. Is that the kind of question you were asking, or were you asking a deeper question?

DR. MURPHY:

6 parts in 10^{12}, for instance, on this particular oscillator. I was wondering if that was a stability figure or an accuracy figure?

DR. ANDERLE:

It is a figure corresponding to the Allan variance.

DR. MURPHY:

Right.

TRANSIT SATELLITE SYSTEM TIMING CAPABILITIES

Theodore D. Finsod
Navy Astronautics Group
Point Mugu, California

ABSTRACT

Within the last year a series of events have occurred with respect to the operational Navy Navigation Satellite System (NNSS), also known commercially as TRANSIT, which may have direct benefits for the Precise Time and Time Interval (PTTI) community.

> In September of 1977 the Director, Strategic Systems Projects granted permission for the Navy Astronautics Group (NAVASTROGRU), in cooperation with the U. S. Naval Observatory (NAVOBSY), to define changes in equipment and operational procedures which could improve TRANSIT satellite time accuracy with respect to UTC (NAVOBSY Master Clock).
>
> For the first time in the 14-year service history of the TRANSIT System, requirements specific to timing accuracy will soon be implemented.
>
> Recent development of a TRANSIT Model T-200 Satellite Timing Receiver by Satellite Navigation Systems, Inc. expedites TRANSIT time transfer capabilities as well as providing NAVASTROGRU an independent measurement of satellite time.
>
> A new generation of superior satellites (NOVA) will be launched starting in early 1980 to augment the existing operational TRANSIT constellation.

This paper reviews the TRANSIT Satellite System in terms of its current time transfer capabilities. Potential improvements using current operational satellites are discussed in terms of changes in equipment and operational procedures which can be made with minimal effort or expenditure and without degradation to the NAVASTROGRU primary (navigational) mission.

INTRODUCTION

The Navy Navigation Satellite System (NNSS) is a fully operational navigation system that enables the Navy Fleet or commercial users to accurately obtain their position anywhere on Earth, day or night, and in any weather. The NNSS, commonly known as TRANSIT, became fully operational in January 1964 and was released for uncontrolled public use by a Presidential directive in July 1967. TRANSIT is operated by the Navy Astronautics Group (NAVASTROGRU), located at Point Mugu, California.

The Applied Physics Laboratory of the Johns Hopkins University (APL/JHU) has played the central role in the technical development of the TRANSIT System. The original idea was conceived there, most of the actual development was performed there, and APL continues to provide technical support in maintaining and improving the system. Figure 1 illustrates the position of NAVASTROGRU in the Navy chain of command.

Today TRANSIT consists of a constellation of five operational satellites in fixed circular polar orbits at an altitude of approximately 1100 kilometers. The characteristics of the current operational TRANSIT satellites as well as their (less than optimal) longitudinal coverage as of 1 April 1978, are summarized in figure 2. All points on the surface of the Earth periodically pass under each orbital path with a nominal time between passes of 90 minutes.

The TRANSIT System requires accurate time to accomplish its navigational mission. For this reason the satellites transmit a precisely timed fiducial time mark (FTM) every two minutes. The timekeeping ability of the operational NNSS is superior to that required for the navigational mission, but is perhaps only marginal in terms of modern Precise Time and Time Interval (PTTI) standards. Within the last year, however, a series of events have occurred with respect to the operational TRANSIT System which may have direct benefits for the PTTI community.

1. In September 1977 the Director, Strategic Systems Projects (DIRSSP) granted permission for NAVASTROGRU, in cooperation with the U. S. Naval Observatory (NAVOBSY), to define changes in equipment and operational procedures which could improve TRANSIT satellite time accuracy with respect to UTC (NAVOBSY Master Clock).

2. The recent development and availability of a TRANSIT T-200 Satellite Timing Receiver by Satellite Navigation Systems, Inc. expedities TRANSIT time transfer capabilities as well as providing NAVASTROGRU an independent measurement of satellite time.

3. In November 1977 the DIRSSP approved the NAVASTROGRU recommendation to install and evaluate a TRANSIT T-200 Satellite Timing Receiver at its Headquarters Computer Center.

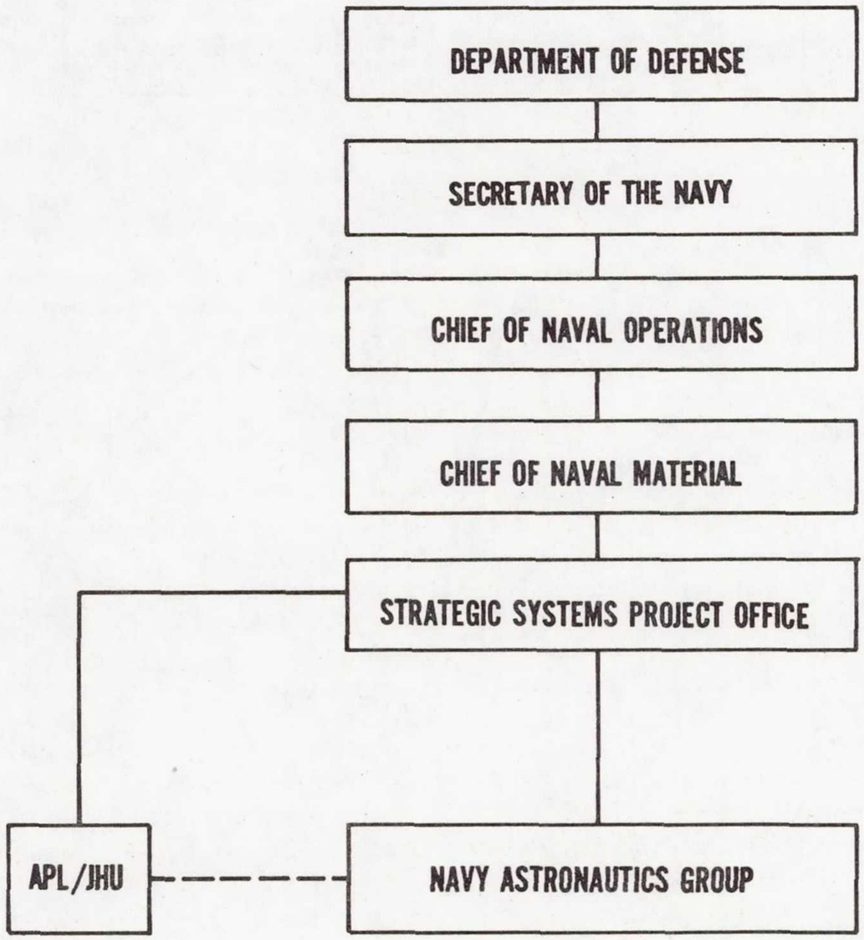


Figure 1. Navy Chain of Command.

CHARACTERISTICS OF OPERATIONAL TRANSIT SATELLITES (APRIL 1978)

SCIENTIFIC NAME	NNSS DESIGNATION	LAUNCH DATE	PERIOD (MIN)	ALTITUDE PERIGEE (KM)	ECCENTRICITY	INCLINATION (DEG)
1967-34A	30120	14 APR 67	106.45	1043.97	.00247	90.25
1967-48A	30130	18 MAY 67	106.94	1071.40	.00179	89.61
1967-92A	30140	25 SEP 67	106.73	1046.58	.00383	89.24
1970-67A	30190	27 AUG 70	106.96	956.96	.01725	90.13
1973-81A	30200	29 OCT 73	105.55	901.95	.01594	90.11

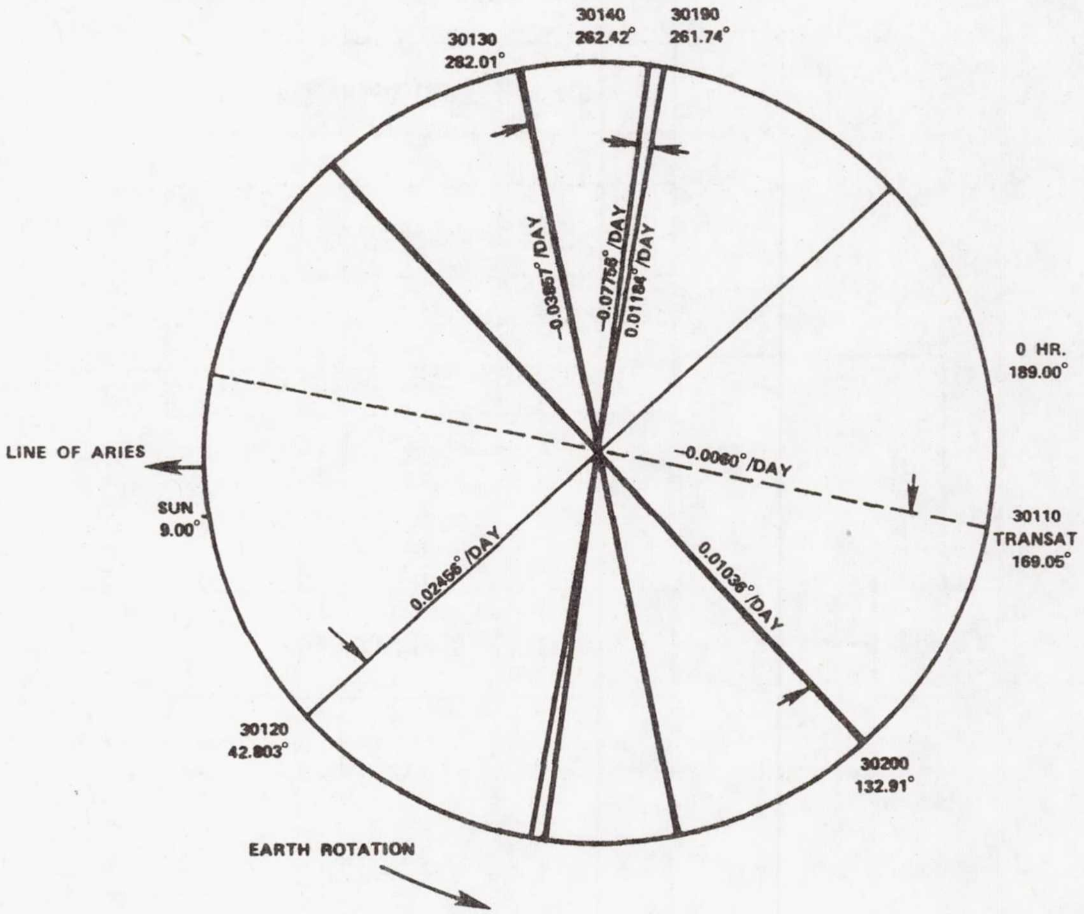


Figure 2. Current Operational TRANSIT Satellite Characteristics.

4. In September 1978 the NAVASTROGRU and NAVOBSY commenced a cooperative test and evaluation program to identify error sources and operational constraints, if any, in transferring TRANSIT time with the TRANSIT T-200 Satellite Timing Receiver.

This report reviews the TRANSIT Satellite System in terms of current time transfer capabilities. Emphasis is placed on time recovery and operational time management. Potential improvements using current operational satellites are discussed in terms of changes in equipment and operational procedures which can be made with minimal effort or expenditure and without degradation to the NAVASTROGRU primary (navigational) mission.

THE TRANSIT SATELLITE SYSTEM

After a brief description of NNSS configuration, we examine in turn the major constituent parts insofar as they determine the time and frequency control of the system.

The NNSS consists of an operational constellation of five TRANSIT satellites and their associated ground support system. The ground support system consists of four tracking stations and a central computer facility. As shown in figure 3, the stations located at Laguna Peak, Point Mugu, California; Rosemount, Minnesota; and Prospect Harbor, Maine, function as both tracking and (message) injection facilities (TRAINFAC's), while the station in Wahiawa, Hawaii, operates only as a tracking facility (TRAFAC). All ground stations have a communication link with the Headquarters Computer Center, located at Point Mugu, California. The operational configuration of the NNSS shown in figure 4 illustrates the salient features which we now summarize.

Each TRANSIT satellite continuously transmits its present ephemeris encoded by phase modulation on two stable carrier frequencies of approximately 150 and 400 MHz. This navigational information is broadcast in two-minute intervals which begin and end at the instant of each even minute. An encoded time marker is part of this broadcast with time uniquely marked at the instant of the even minute. All broadcast frequencies, as well as the satellite clock are based on a highly precise master oscillator.

As pass geometry and operational considerations permit, NAVASTROGRU stations track each satellite to obtain doppler information and the raw timing data. After the satellite has set (typically, 17 minutes elapse from rise to set), the tracking data are transmitted to the Headquarters Computer Center where all measurements from all tracking stations for each satellite are accumulated. At least once a day the data are used in a large computing program to:

1. Determine the contemporary orbit specification for the satellite and predict an ephemeris for the next 16 hours.

2. Compute the necessary corrections to the satellite clock to compensate for the predictable part of the oscillator drift.

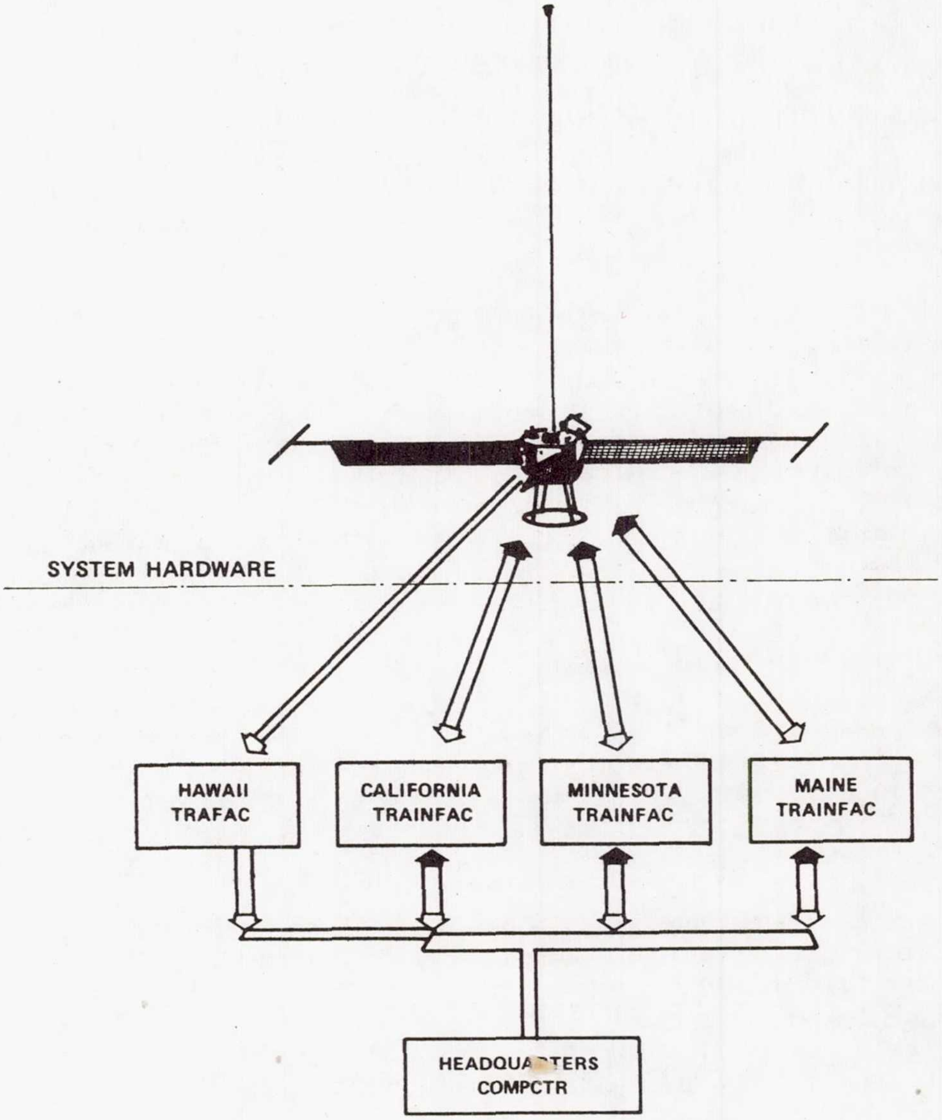


Figure 3. TRANSIT Ground Support System.

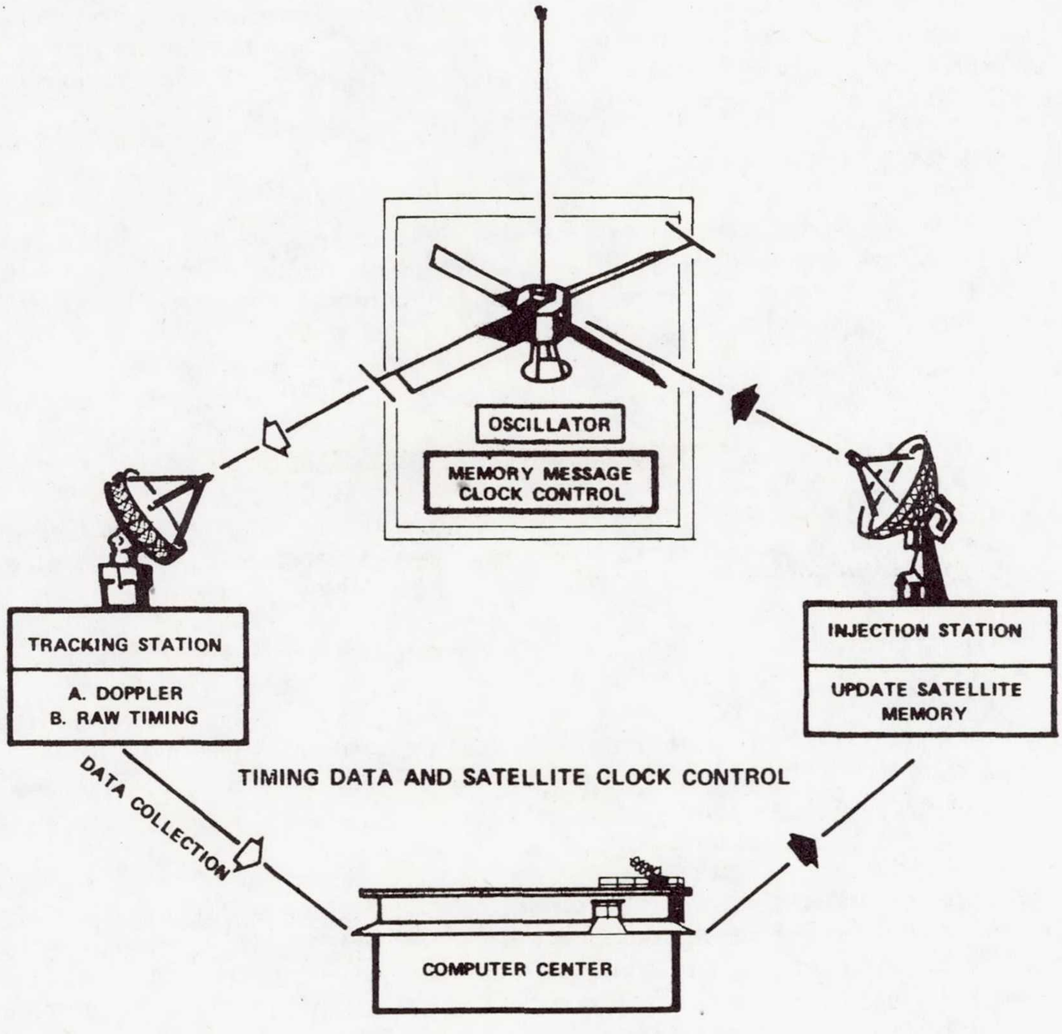


Figure 4. Operational Configuration of the NNSS.

The ephemeris prediction and satellite clock correction information is then transmitted back to all three injection sites. One station performs the injection; however, the mutually-in-view station provides backup in the event of equipment failure. The injection station inserts an updated ephemeris into the satellite memory. Typically, injections into the satellite occur at 12-hour intervals even though the satellite memory has sufficient storage to contain a 16-hour ephemeris. To navigate using a TRANSIT satellite requires a special tracking receiver/computer. This receiver measures the broadcast carrier frequencies at discrete intervals to recover doppler information and simultaneously demodulates them to recover the satellite ephemerides. A small computer is necessary to then calculate the position of the navigator.

TRANSIT SATELLITE CLOCK

The TRANSIT satellite contains an ultrastable 5-MHz oscillator. This oscillator is a temperature-stabilized, crystal-controlled transistor oscillator designed to generate an extremely precise frequency. It is used as the standard frequency source for all satellite operations.

The exact operating frequency of the oscillator will, of course, vary slightly with circuit age and environment. The nominal value or design center is specified as 4,999,600.0 Hz, an offset of 80 parts per million below 5 MHz. The actual value is expected to remain within a tolerance of ± 10.0 Hz (2 parts in 10^6) over the entire operational life of the satellite. Frequency drift rates in flight have been on the order of several parts in 10^{11} per day. Over the time of a single pass (approximately 20 minutes) the drift rate has generally been several parts in 10^{12}.

TRANSIT satellites transmit coherent carrier frequencies of 150 and 400 MHz (offset -80 ppm) by multiplication of the reference oscillator frequency and hence share the same stability. The satellite navigation message will be transmitted on these carriers in the form of a *double-doublet* pattern of phase modulation. The memory read rate is controlled by a clock consisting of a divider chain reducing the master oscillator frequency.

The satellite clock is constructed as an integral part of the memory readout system. Operating from timing provided by the master oscillator, this system is designed to read a complete message periodically every two minutes. A timing mark is included near the beginning of each readout in order to generate a well-defined two-minute interval. Fine control over the timing (or epoch) of the clock is provided so accurate synchronization can be maintained with UTC.

Satellite timing is synonymous with memory readout rate, and timing control is contained within memory content. We therefore describe TRANSIT memory organization before discussing the control of the FTM.

The memory of the TRANSIT satellite consists of 640 thirty-nine-bit words which are divided into a main or a *fixed* memory of 160 words and a variable or an *ephemeral* memory of 480 words. Figure 5 illustrates the TRANSIT satellite memory message format, identification words, and location of the FTM. Although main memory contains 160 thirty-nine-bit words, the satellite will broadcast 6103 bits (i.e., 156 words plus 19 extra bits) evenly spaced over each two-minute interval. (The remaining bits are used for real-time operational commands.) To recover timing and message information from a TRANSIT satellite broadcast a ground receiver must identify the beginning of a message transmission and synchronize with a unique bit pattern to determine the FTM. The first three words of each message broadcast cycle accomplish these goals.

The first word is a BARKER word. The first 33 bits of this word are the barker pattern and are arranged in such a fashion that they never can be repeated in any other part of the message. This allows a ground receiver to identify the starting point of the message. This 33-bit pattern is the same for all TRANSIT satellites. Bits 34 through 37 of this word form a unique satellite identification code while bits 38 and 39 are always zero.

The second word is the synchronization word or SYNC word. The first 14 bits are special-purpose bits. The 15th bit is a zero followed by 23 one's and another zero. It is this final 25-bit pattern which is used in conjunction with the BARKER word by the NAVASTROGRU navigation receivers (AN/BRN-3) to determine the FTM, although by definition the FTM is the first phase transition of the first bit of the third word. This third word is called a BEEP word because the unique bit pattern associated with it produces an audible *beep* of approximately 400 Hz in NAVASTROGRU receiving equipment.

The remaining 153 words of the readout cycle in general contain the information comprising the navigation message to be broadcast. This message (which is periodically transmitted from the ground and stored in the satellite memory) consists primarily of a set of Kepler parameters specifying the current orbit of the satellite. Fine corrections to these *fixed* parameters (which are used by the navigator on the ground to compute a more accurate current satellite position) are provided in a set of eight special words, called ephemeral words. These words taken in pairs contain the corrections needed for a particular two-minute interval.

As indicated in figure 5, each ephemeral word in the set of eight is furnished each readout interval. To keep the set referenced to the indicated past, present, and future two-minute intervals, the set is shifted down six memory locations each readout cycle. The most future entry of the set (word 50) enters from ephemeral memory (word 161) while the oldest entry (word 8) is lost. Since one new word is required each two-minute interval and ephemeral memory can furnish 480 words, the satellite can operate for 16 hours before exhausting its current orbit information. (Injections of updated orbit information are made on a 12-hour basis.)

SATELLITE MESSAGE FORMAT

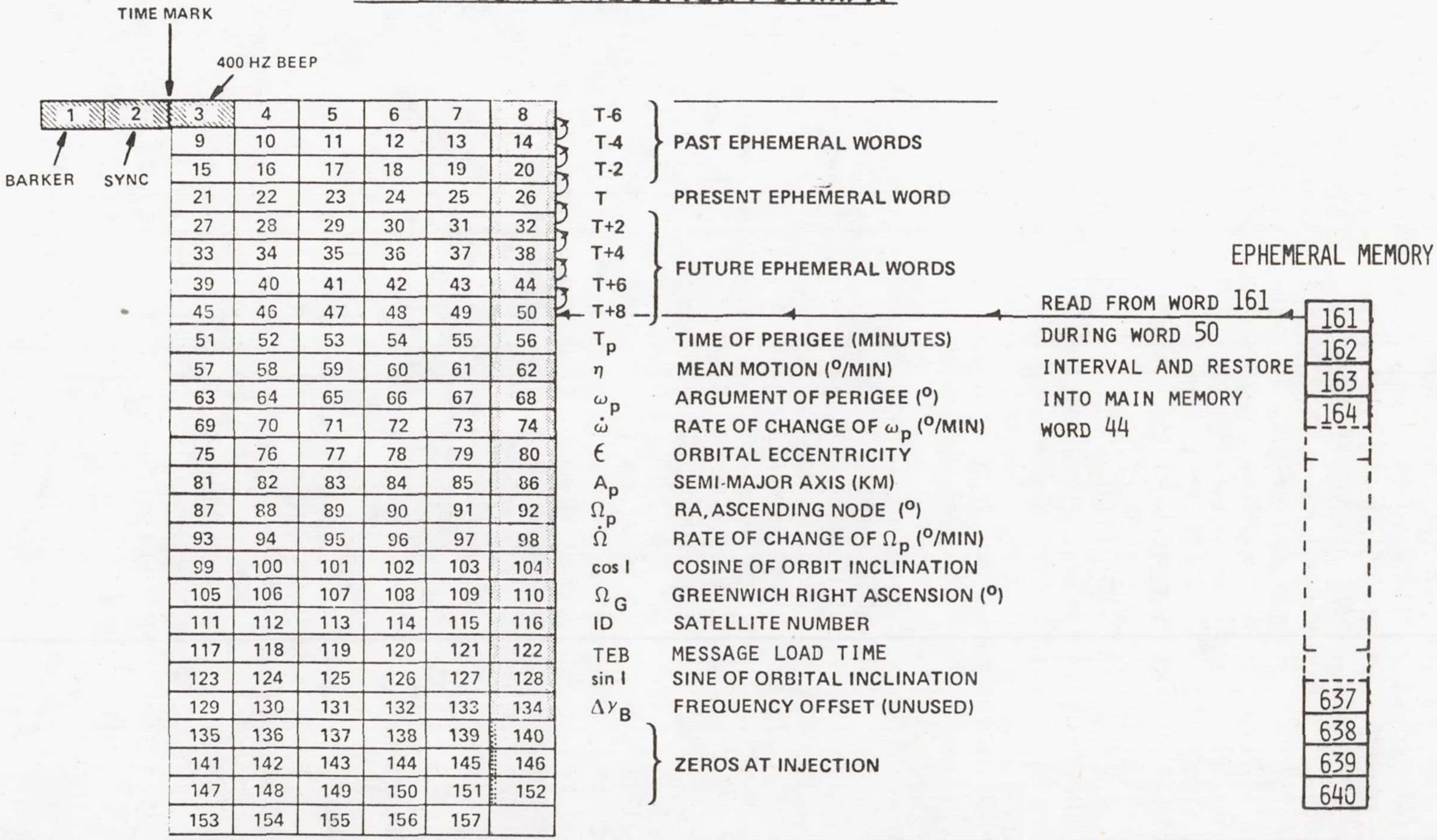


Figure 5. Satellite Message Format.

In addition to the words used for the navigation message, there are special words and bits which are used for certain satellite monitoring and control functions. The last three bits of all memory words fall into this category. Only the first 36 bits of each 39-bit word are used for message data. The last three bits of each have special functions; they are designated P, T, and C, for parity, telemetry, and clock correction, respectively. Of interest here are the clock correction bits, which are used to make fine corrections in the satellite timing. The 39th bit of most main memory words through word 136 and all ephemeral words is available as a clock correction bit.

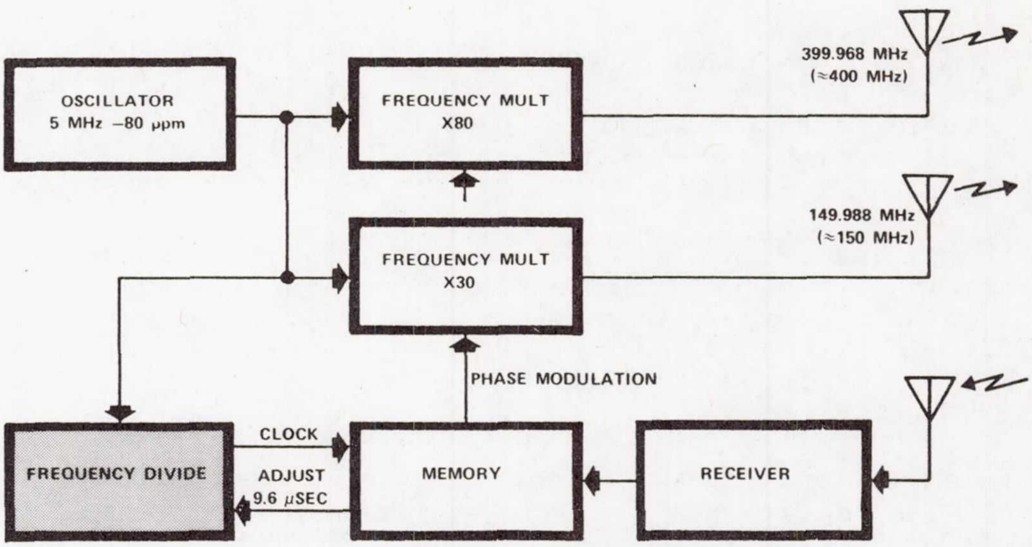
The discussion of satellite hardware which allows these clock correction bits to function as a vernier time control is aided by reference to figure 6. At the top of this figure is the general block diagram illustrating the general features of TRANSIT signal generation previously discussed. We concentrate now on the divider chain (FREQ DIVIDE block) between the master oscillator and the MEMORY block.

This divider chain is constructed with sequential divisions of 16, 3, 2, 32, and 32, yielding a total frequency division of the master oscillator of 98,304. The memory read clock thus formed has a frequency of 50.859 Hz corresponding to a read period for each bit of 19.662373 milliseconds. Since the total message consists of 6103 bits, the transmission time is 119.9994624 seconds, which is 537.6 microseconds short of a precise two-minute interval. The adjustment of the memory read clock to a precise two-minute interval is made with the clock correction bits. Figure 6 shows this link is made via the ADJUST 9.6 μSEC line from the MEMORY block to the FREQ DIVIDE block and constitutes the vernier time control.

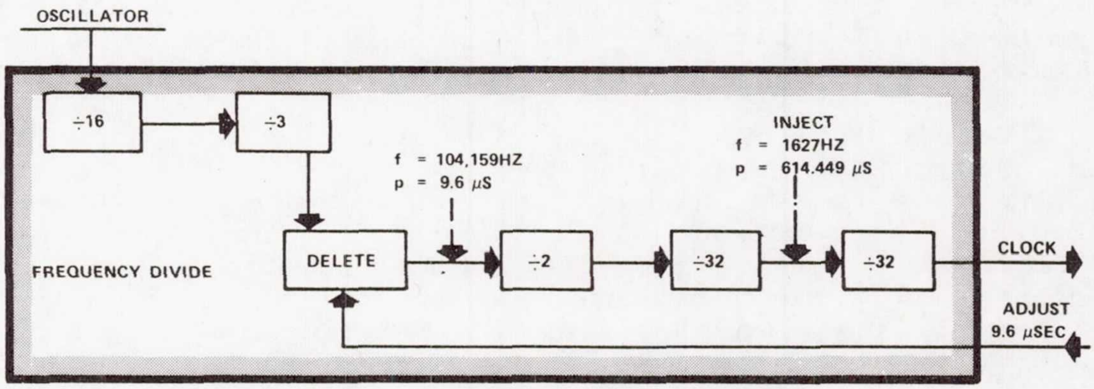
The vernier time control system, called the *time normalizer*, operates as follows. Whenever a word is read from main memory, the 39th bit is checked for content. If it is a zero bit, satellite timing continues uninterrupted. However, if it is a one bit, a pulse activates the DELETE circuit. This circuit simply delays by exactly one period the signal at its input. This input signal frequency having been reduced from the master oscillator by a factor of 48, has a period 9.60 microseconds. This action delays all memory readout operations lengthening the duration of the bit, and hence the entire readout cycle by that amount.

Since each clock correction bit (if a one bit) lengthens the interval of the message readout by 9.6 microseconds, to achieve an exact two-minute interval 537.6 microseconds ÷ 9.6 microseconds = 56 clock correction bits are required. This, of course, assumes a nominal master oscillator frequency of 4,999,600.0 Hz. To provide both positive and negative compensation about this nominal operating point, an additional 56 clock correction bits are required. This allows a time adjustment of ± 538 microseconds per two-minute interval. In fact, of the possible 156 words in main memory exactly 124 words are available (the remaining 32 are not used for various reasons). Since the change of a single time control bit in main memory results in a cumulative displacement of the FTM by 9.6 microseconds *every* two-minute period, over a period of 12 hours this amounts to a net displacement of the FTM of 3.456 milliseconds.

TRANSIT SIGNAL GENERATION



SATELLITE CLOCK CONTROL SYSTEM



1. SATELLITE OSCILLATOR DESIGN CENTER FREQUENCY = 4,999,600
2. DIVIDER CHAIN RATIO (READ) = 16 x 3 x 2 x 32 x 32 = 98,304
3. 98,304/4,999,600 = 19.662373 ms/BIT
4. 19.662373 x 6103 = 119.9994626 SECONDS
5. 120.0000000 − 119.9994624 = 537.6 μS
6. DIVIDER CHAIN RATIO (DELETE) = 48/4999600 = 9.6 μS
7. 537.6 μS/9.6 μS = 56 CLOCK CORRECTION BITS/2 MIN. INTERVALS

Figure 6. Satellite Clock Control System.

Finer control than that which can be obtained by inserting (or deleting) a full control bit each two-minute cycle is available in the form of clock correction bits stored in the ephemeral memory words. These bits are used only once, when they are read from the ephemeral memory at word 50, and thus they provide a unique correction for each individual two-minute interval. Since each bit is used only once in the entire normal 12-hour span between injections (which reload the memory), it has one three-hundred-sixtieth the effectiveness of the time correction bits in the main memory. *With proper use of these bits, together with those in main memory, it is theoretically possible to hold the clock on time to an accuracy of ± 4.8 microseconds over any 12-hour period.*

The operation of the *time normalizer* has been described in terms of time increments, but in practice its principal function will be to compensate for master oscillator frequency variations in such a way as to hold the length of the readout interval to exactly two minutes. When the oscillator speeds up, the proper number of additional one bits will be inserted as time correction bits to slow down the readout and maintain a constant cycle length. When the oscillator slows down corresponding bits will be removed to match. In this way, the clock can be held to a fixed two-minute cycle in synchronization with UTC. The amount of frequency compensation corresponding to a single clock correction bit is given by $48 \div 120 = 0.4$ Hz, since each bit controls a deletion of 48 cycles from the total two-minute interval. The total amount of control available, corresponding to ± 56 clock correction bits, is a frequency compensation of ± 22.4 Hz. It can be seen that this control is more than enough to cover the expected lifetime drift (± 10 Hz) of the oscillator with some left over for time synchronizing operations.

This, then, is the satellite clock system -- a stable oscillator driving the memory readout through a frequency divider chain. Synchronization of the FTM readout with UTC is achieved through control of the frequency divider chain using special (clock correction) bits embedded in the satellite message. Thus synchronization is controlled from the ground through the use of clock correction bits transmitted and stored in the satellite memory. In operation the system should consititute a precision clock which can be accurately synchronized with UTC.

To understand how TRANSIT time is maintained; i.e., how the proper number of clock correction bits is determined, we proceed to an operational description of the ground control system.

THE GROUND CONTROL SYSTEM

The ground control system consists of a tracking facility (TRAFAC), tracking and injection facilities (TRAINFAC's), and the Headquarters Computer Center. Figure 7 is a functional block diagram of a TRAFAC/TRAINFAC and the Headquarters Computer Center.

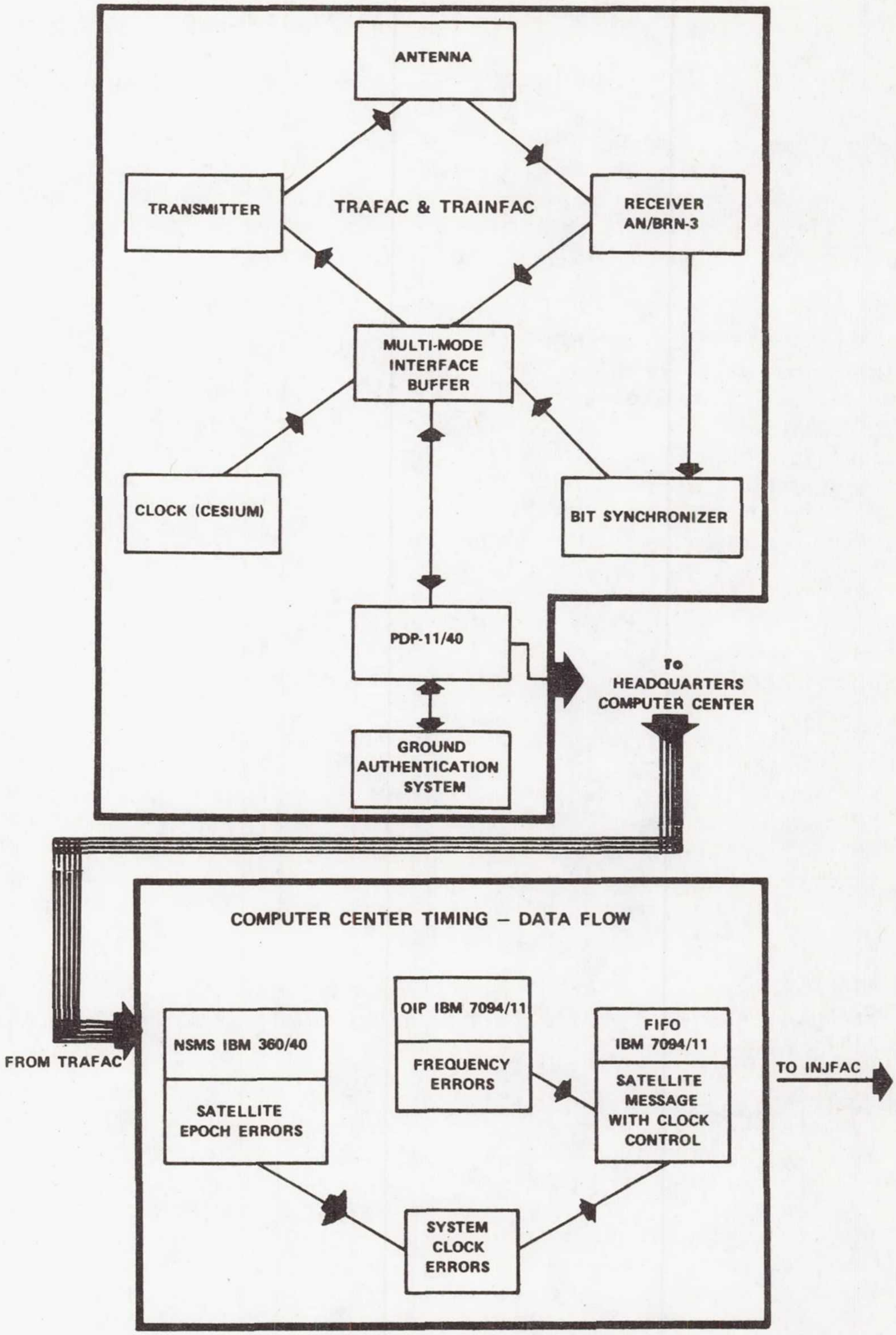


Figure 7. Functional Diagram of Ground System Facilities.

The principal function of the ground control system is to compute the proper pattern of time correction bits required to hold the clock on time with respect to the chosen referenced time, UTC. This calculation can be thought of as being carried out in two steps (although the results of both steps are, in practice, combined to form a single pattern of bits to be included in the injection message and transmitted to the satellite). First, the proper number of correction bits required to hold the desired *clock rate* is determined according to the current oscillator frequency. Second, the number of bits needed to correct any error in the *clock epoch* is calculated in such a way that the error will be removed over a normal 12-hour injection period.

Doppler and timing information collected by the TRAFAC from a TRANSIT satellite pass is sent to the Headquarters Computer Center. The timing data are processed in near real time by the NAVASTROGRU Satellite Monitoring System (NSMS) programs. In this program raw timing data consisting of received FTM's relative to the TRAFAC Cesium clock are processed to obtain the FTM epoch referenced to the satellite and relative to UTC. The output of the NSMS is a plot of the time errors associated with passes over a time span of the previous 36 hours. From this plot, a judgment is made by an analyst as to the clock epoch error. An example of this plot is shown in figure 8. The expected epoch error, in this case 30 microseconds, will be sent to the Final Formatting (FIFO) program. In addition, a second input to FIFO is the result of the doppler data analysis. From this analysis, the master oscillator frequency and drift rate are computed at a predetermined epoch and linearly extrapolated over the injection message readout span. (The program providing this analysis, the Orbital Improvement Program (OIP), also supplies the updated ephemeris for the next injection span.)

With the inputs of satellite clock epoch error and oscillator drift rate, FIFO calculates and inserts the proper number of clock correction bits into the injection message. Subsequently, these injection messages are transmitted to the TRAINFAC for injection in the proper satellite at the proper time.

CURRENT TRANSIT TIME ACCURACY

Having delineated how the NNSS generates and maintains TRANSIT satellite time, consideration will be given to the limitations imposed by the hardware and the current time dissemination accuracy of the TRANSIT satellites.

Although the inherent limit to timing precision is that of the satellite clock (9.6 microseconds), this limit is rarely an impediment in the current operational system. All TRAINFAC's use an AN/BRN-3 navigation receiver which can recover a single FTM with a precision of approximately 35 microseconds (1σ). The signal delay time from the AN/BRN-3 receiver, although measured weekly, is a function of many variables, and the correction uncertainty is approximately 15 microseconds (1σ). In the tracking of a single pass an average of six FTM datum points is recovered, thus improving the time precision by a factor of $\sqrt{6}$. Each TRAINFAC is equipped with a cesium clock whose time with respect to UTC is maintained (via traveling clock checks) within five microseconds (1σ). Thus, the expected precision in a time measurement of a single pass of a TRANSIT satellite within the NNSS is 19 microseconds (1σ).

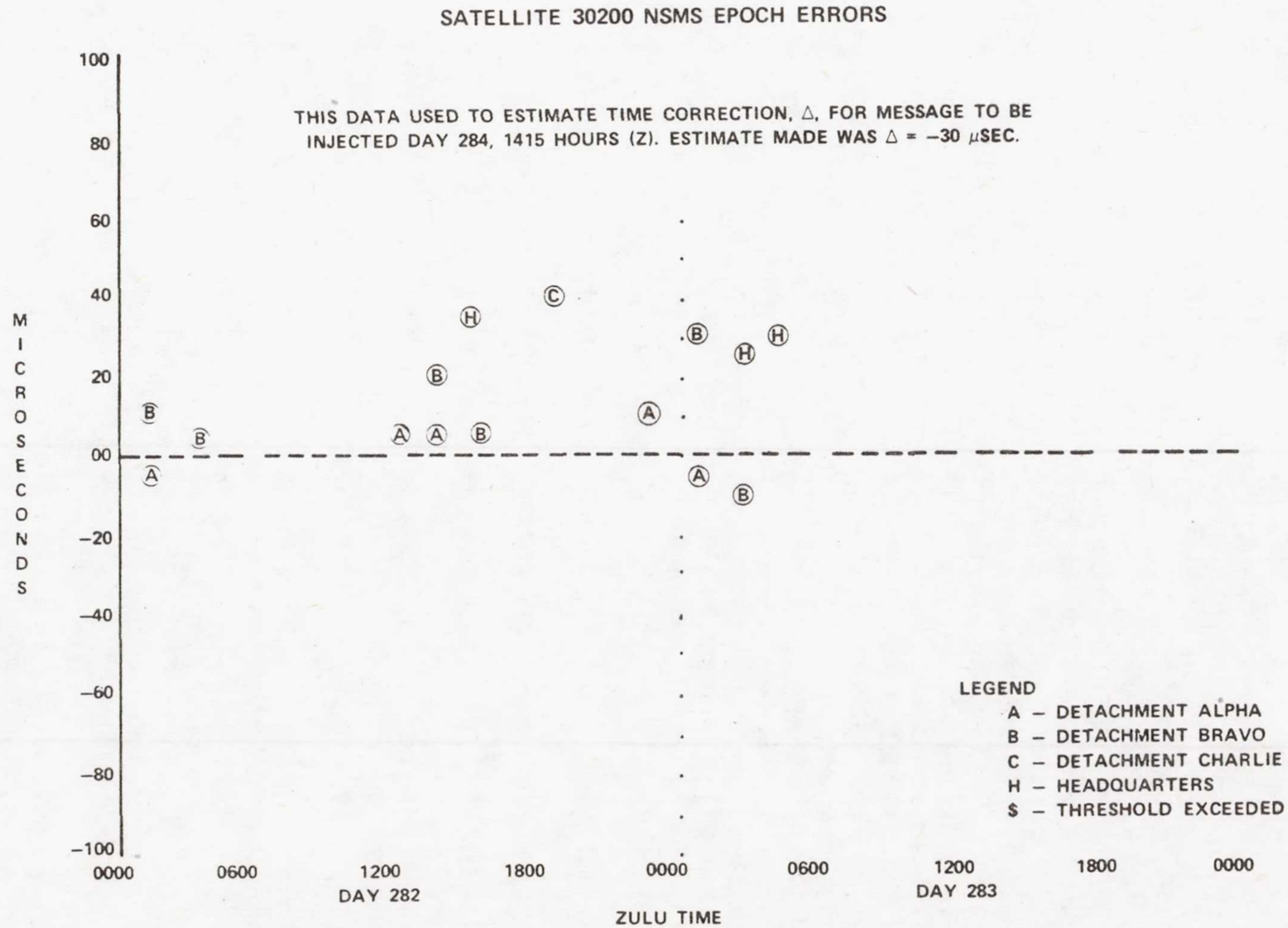


Figure 8. Satellite 30200 NSMS Epoch Errors.

With this basic hardware limitation of measurement precision, figures 9 and 10 exemplify the current accuracy of each TRANSIT satellite referenced to UTC.

Immediately apparent is the disparity in time accuracy between the older and newer satellites. The data base for each satellite consisted of over 1200 passes averaged over all tracking stations. Special problems with the two oldest satellites are understood: satellite 30120 has reduced power on the transmission of the 400-MHz carrier; and on satellite 30130 the 150-MHz tranmission has the satellite telemetry signal (2.3 kHz) superimposed on it. The single-pass accuracy of the system (all satellites combined) is approximately 19 microseconds (1σ).

Another feature apparent in the time accuracy histograms is the relative frequency of satellite clock errors over 70 microseconds. The sources of these errors are satellite oscillator jumps and ground station injection problems. However, these problems are known at the time of their occurrence and corrected within a reasonable time frame. Table I illustrates the NNSS operational time dissemination reliability maintained over the previous four years.

TIME IMPROVEMENT PROGRAM

Improvement in TRANSIT timing accuracy in the near future should be considered in terms of measurability and controlability. Both aspects are limited ultimately by the satellite clock. The most important limitations in practice, however, are delays throughout the feedback loop of the control operation. This limitation becomes critical when compounded by limitations on measurability. When time thresholds are implemented which require quick and serious response, it is imperative that reliable confirmation be obtained as to the time measurement in question.

Feasibility of improved measureability was made possible by the introduction last year of a new TRANSIT timing receiver by Satellite Navigation Systems, Inc. A paper on the Model T-200 Satellite Timing Receiver was presented at the Ninth Annual PTTI conference (1977). Suffice it to say that initial evaluations indicate a time recovery capability comparable to the NSMS. This receiver, if used at each TRAFAC, would give an independent measure of time for each satellite tracked, thereby redundantly doubling the number of datum points per pass. This allows a more reliable determination of satellite timing which ultimately results in better control.

Feasibility of improved controlability would come through NAVASTROGRU operational procedure changes. To illustrate the extent of the delay in the control loop consider figure 8 again. The estimate of the clock epoch error was made as of 1600Z on day 283.

The correction, -30 microseconds, was injected into the satellite memory at 1415Z on day 284. The corresponding clock correction bits were spread as uniformly as possible throughout the 12-hour readout period. This is a delay in correcting a known clock error of 36 hours! Obviously any time improvement program initiated by NAVASTROGRU will place high priority on shortening this control loop delay.

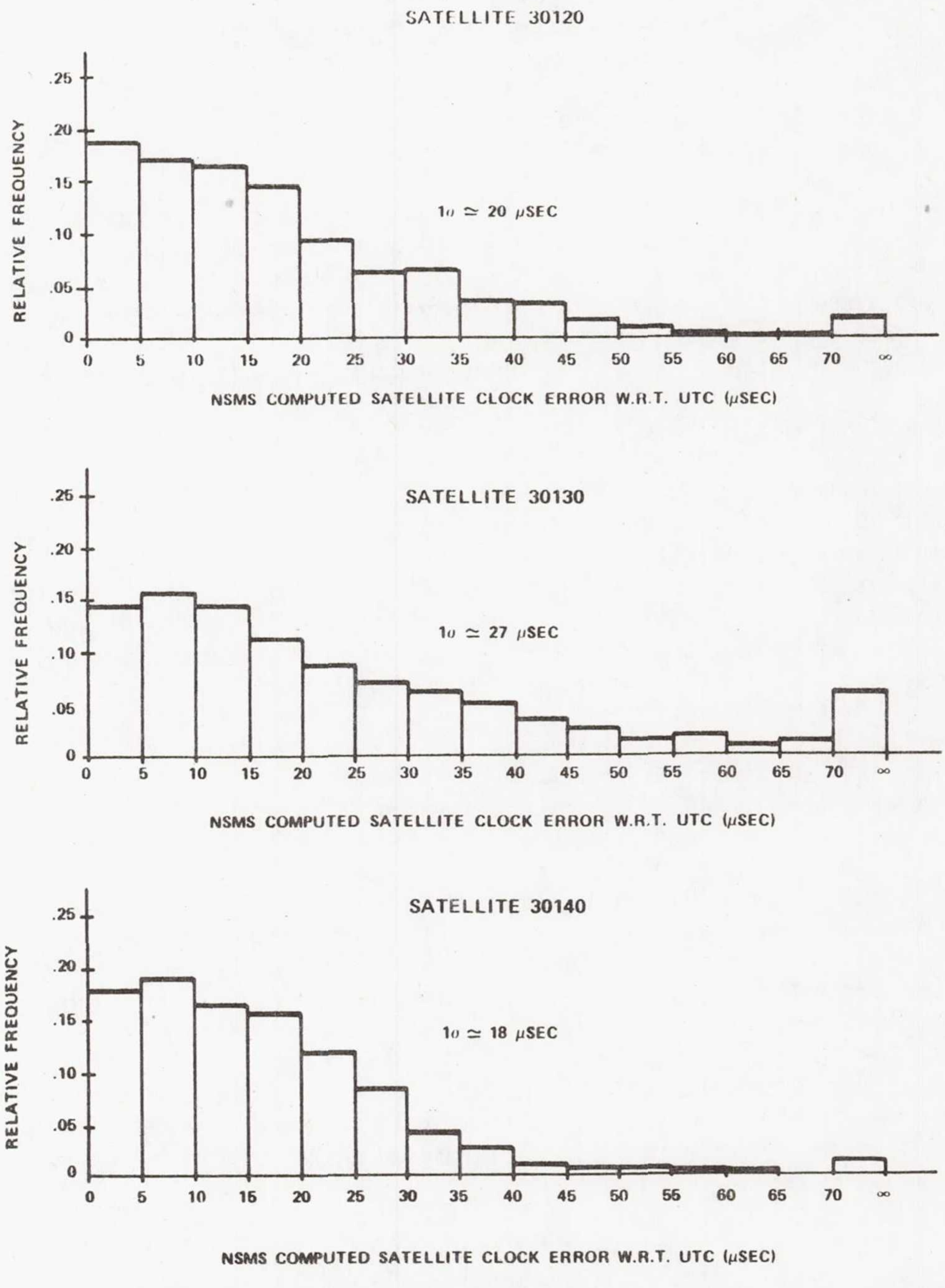


Figure 9. NSMS Computed Satellite Clock Errors with Respect to UTC.

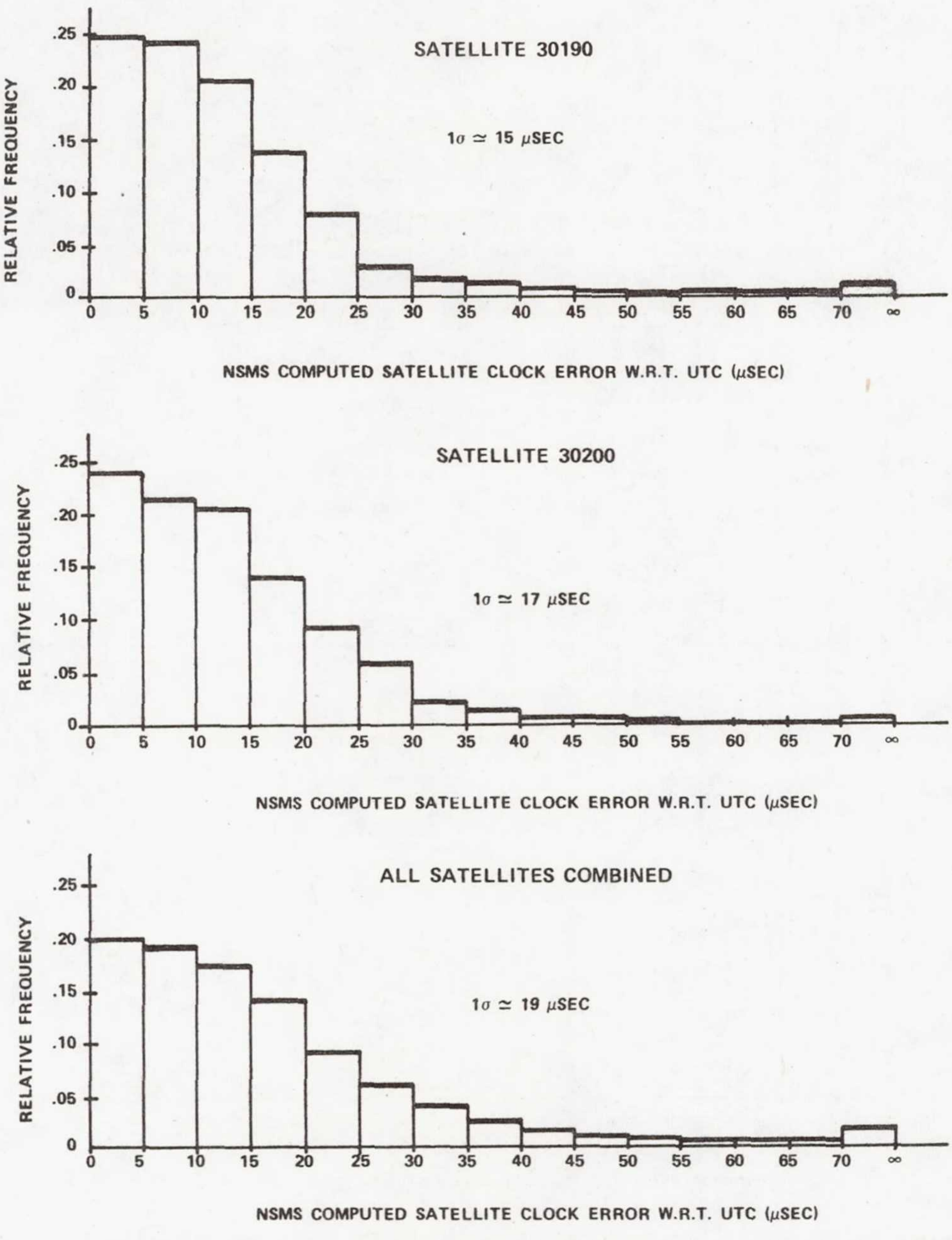


Figure 10. NSMS Computed Satellite Clock Errors with Respect to UTC.

TABLE I. NNSS OPERATIONAL TIME DISSEMINATION RELIABILITY

SATELLITE BROADCAST EPOCH > 100 μSEC	
YEAR	INSTANCES
1978	5
1977	8
1976	14
1975	7
1974	4
CATEGORIES OF SATELLITE TIMING ANOMALIES	
- DETACHMENT INJECTION UPLINK	33
- SATELLITE HARDWARE MALFUNCTION	4
- OIP FREQUENCY MISPREDICTION	1
RELIABILITY	
- PERIOD SINCE 1 JANUARY 1974	
- APPROX 233,000 BROADCAST HOURS	
- APPROX 1088 HOURS SAT CLOCK > 100 μSEC	
- RELIABILITY = 99.5%	

The following hardware-related modifications are expected, either directly or indirectly, to improve the accuracy of time disseminated by the TRANSIT satellite system.

1. If, in the current joint evaluation (by NAVOBSY and NAVASTROGRU) the TRANSIT timing receiver performs as well as initial tests indicate, NAVASTROGRU will request (of DIRSSP) deployment of one receiver to each detachment.

2. Continual upgrading of supplementary helix tracking antennas at all TRAINFAC's for improved reception of satellite signals. This is especially important as we enter a period of high solar activity.

3. Deployment of a second cesium clock to the detachments at Maine, Minnesota, and California to improve the quasi-yearly checks by NAVOBSY.

4. The use of TV Line-10 from Los Angeles KTTV (Channel 11) by Headquarters as a continual link to UTC by the TRANSIT System. (As of November, 1978, this TV Line-10 source is still not operational.)

Software-related improvements are under consideration but are more tenuous due to cost factors. They would consist mainly of improvements to the NSMS timing program.

In summary, the operational TRANSIT System meets the current NNSS timing objectives. NAVASTROGRU feels that with the introduction of the hardware and operational changes mentioned above, anticipated NNSS requirements specific to timing accuracy can also be met. Any further requirement, however, for time transfer to a higher degree of precision on a continuous basis must be established through the Chief of Naval Operations.

QUESTIONS AND ANSWERS

MR. LAUREN J. RUEGER, Johns Hopkins University, Applied Physics Lab:

I feel impelled to defend the AN/BRN-3 Receiver design, as I had a part in its development. It was designed to meet a specification of 100 microsecond time recovery in 1969 and I am a bit surprised that it exceeded that by as large a margin as you have shown. You did not mention the specification of the navigation system for timing at the present time. Do you want to mention that?

DR. FINSOD:

Yes. There is no timing specification now. Our primary mission, which is navigation, determines that. In those terms, people don't get excited until they see 1,000 microseconds deviation; then the duty officers, or the controllers, get excited. I call them up when the deviation is about 200 microseconds and I start recovery action at around 20 microseconds. So, because there is no specific requirement, it is quite a mess in determining who gets excited when, but we do definitely make changes. When there is a 20 microsecond deviation, we do put in delts, what we call delts, or corrections.

MR. RUEGER:

I thought there was in the works some requirement to hold 200 microseconds relative to the Naval Observatory.

DR. FINSOD:

Right.

MR. RUEGER:

Is that not in effect yet?

DR. FINSOD:

It is public knowledge that the requirements that we are talking about, which are coming downstream soon, are to maintain no more than a 70 microsecond difference in any satellite clock on an orbit-to-orbit return. In other words, you go around twice and you can't deviate more than 70 microseconds, nor can you allow any more than 200 microseconds at any time on any satellite.

DR. ALFRED KAHAN, Rome Air Development Center:

Can you comment about the physical performance of the oscillators for those last ten years. For example, have you observed any solar activity on those oscillators?

DR. FINSOD:

If there has been, I don't believe that we have seen it. On the older satellites, we have had oscillator jumps, but I think that is completely unrelated. We have found no timing problems related to

solar activity, except those caused by increased scintillation of the ionosphere, which causes problems in reading time. But we have seen no evidence, that is, they are maintaining their offset of 80 parts per million within something like 50 parts per 300 million. I mean, that is sort of the worst offset we have to the requirement, and this is after 11 years. There has not been any correlation that we have seen.

I think Mr. Rueger would like to comment.

MR. RUEGER:

We do have data on that subject and we observed the natural radiation effects on the crystal oscillators which I suspect you are seeing. The normal aging drift in these oscillators is up in frequency, but down in frequency from the radiation effects of trapped electrons in orbit. The upward drift is linear with time; the downward drift from radiation dose is logarithmic with time. So, all these satellites drifted slightly down, at first, and then they eventually are dominated by the linear function and drift upward. These oscillators have had such a large dose by now that even fairly substantial solar flares cause almost no effects.

We did observe, when we had some very large solar flares on occasion, a jump of the frequency during the early life in orbit by as much as 4 parts in 10^9. We have records of the frequency as a function of time which you can plot yourself from the Naval Observatory Bulletin 17 literature. Those numbers come from the Naval Astronautics Group and are then reissued by USNO to the public through the Bulletin 17.

MR. FRAUHOFF, Efratom California:

With the clock specs that you have, what is the one sigma navigation solution that the man comes up with? Did I mishear that or had that been mentioned?

DR. FINSOD:

I am not sure I quite understand. It sounds like a question that might be directed to our first speaker. Are you talking about the corresponding navigational error?

MR. FRAUHOFF:

Yes.

DR. CHARLES MARTIN, Defense Mapping Agency:

It is based on a single pass.

DR. FINSOD:

Based on a single pass. I don't know that figure off hand. Bob Payne might be able to add something to that.

MR. ROBERT PAYNE, Naval Astronautics Group:

As Dr. Finsod pointed out, our timing is controlled within the auspices of our navigation today and the navigation resolution is on the order of about 3/100 of a nautical mile, about 52 meters, that we can maintain very well over any 24 hour arc. Now, we have been experiencing increased solar activity, and, as you gentlemen know, it does have quite an effect on navigation. We have seen problems that have increased that number by a magnitude of 2 over what I said, so .06 or .07 nautical miles is not unreasonable for very high K or very high values of S average.

SPEAKER:

This comment, I think, gets closer to the answer that he asked.

DR. FINSOD:

Yes, I understood the question was about the clock performance on navigation accuracy and the number I was giving was something of the order of 10 centimeters to a meter, depending on what oscillator you are dealing with. I think the errors you are speaking about are mainly from ephemerous errors rather than from oscillator errors. That perhaps was not what he was asking.

MR. DOUGLAS TENNANT, GPS Program Office:

I am curious to know if you have experienced anything like discreet frequency shifts in the clocks?

DR. FINSOD:

Yes. In fact, a number within the last three or four months, particularly, I think in Satellite 13, or 30130, but we have found no correlation between these and any other activity.

MR. TENNANT:

How about solar activity?

DR. FINSOD:

No. I mean, that is, none of the other satellites have the jumps and the jumps do not occur very often--once every six months usually, which causes, of course, a serious timing problem until it can be corrected.

MR. TENNANT:

In GPS, we have seen discrete frequency shifts which we have correlated to substorms on the sun. I wonder if you had a common experience?

DR. FINSOD:

No, although recently we have been trying to find some correlation because our navigation has deteriorated somewhat, although not out of specifications, and we could find no correlation in the K

indices. But apparently there was high auroral activity, high solar proton activity in the northern latitude where our tracking stations are located, especially detachment BRAVO which is in Minnesota, and detachment ALPHA which is in Maine. So, when we have problems, it is hard to separate out the problems caused by from an excited ionosphere and those that correspond to the actual satellite being bombarded by radiation.

So, we haven't been able to determine any correlations.

Page intentionally left blank

TIME SYNCHRONIZATION VIA THE TRANSIT SATELLITE AT MIZUSAWA

T. Hara and K.-H. Sato

International Latitude Observatory of Mizusawa
Mizusawa-shi, Iwate-ken, 023 Japan

ABSTRACT

Time signals emitted from Transit satellites have been received by the NACODE type receiver since 1974 at Mizusawa, Japan (station 027). By using these time signals, we can make a time comparison between the International Latitude Observatory of Mizusawa (ILOM) and USNO. To complete time comparison by this method, many corrections are, however, necessary such as propagation delays, a receiver delay, effects of relative motion of satellites, effects of the ionosphere and so on. Propagation delays are calculated from the precise ephemeris of the satellite (30190) supplied by the Topographic Center of DMA. The receiver delay is measured by supplying a simulated signal to the space near the receiving antenna. Effects of the ionosphere on the propagation delays may be the order of one microsecond. Standard deviations of each pass are estimated to be ± 15.5 microseconds for the data UTC(ILOM)-UTC(USNO) obtained in December 1976.
Time comparisons by the Loran-C system between ILOM and USNO are referred for a check of the Transit satellite timing method.

INTRODUCTION

Timing experiments via satellites have been carried out many times since 1962 (Blair 1974). In Japan also, experiments of time synchronizations between the Radio Research Labolatories (RRL) and the U.S.Naval Observatory (USNO) were carried out in 1965 and 1975 with accuracies of one microsecond and 10 nanosecond's order respectively (Frequency Standard Section and Kashima Branch 1965; Yamamoto et al. 1976). These experiments were made in the two way method and attained to the very high accuracy.

This method is, however, much expensive and is not convenient for the frequent measurements.

Although the measurements via the Transit satellite (Navy Navigation Satellite) have relatively low accuracies as compared with the two way method, this method has the advantage that time comparison can usually be made twice a day at moderate expence. The Navy Navigation Satellites have been tracked by the TRANET I type receiving system since late 1974, and time information data in punched paper tape are available since 1976. In this report, timing analysis and various corrections which are necessary to derive time differences between UTC(ILOM) and UTC(USNO) are presented by using the data obtained in 1976. Time synchronization via Loran-C system will be referred to examine the consistency of these two methods.

OUTLINE OF COMPARISON SYSTEM

Satellite trackings by measuring doppler shifts have been made with the rubidium oscillator as a frequency standard at the station 027. Time and frequency comparisons have been made between the rubidium atomic clock and UTC(ILOM) which is maintained by a cesium atomic clock. At the same time, satellite timing pulses are monitored by UTC(USNO) and the results are published regularly. Then, time differences between UTC(ILOM) and UTC(USNO) can be derived by using these data.

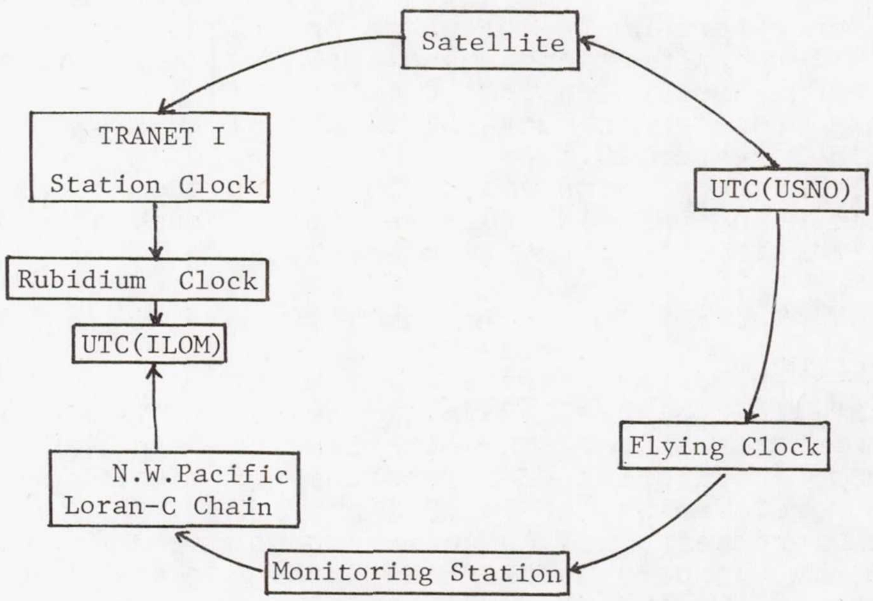


Fig. 1- Simplified block diagram of the time comparison system.

On the other hand, Loran-C signals of the Northwest Pacific chain are received at ILOM regularly with standard deviations of less than 0.1 microseconds. This chain (SS3) is monitored by flying clock from USNO. Thus we have another method of time comparisons between UTC(ILOM) and UTC(USNO). An outline of our time comparison system is shown in Fig. 1.

The TRANET I system receiver amplifies and demodulates signals from satellites, and demodulated signals are fed to the time burst detector which discriminates the satellite time marks. Time interval between this fiducial time mark and the station clock is measured to one microsecond.

CORRECTIONS FOR PROPAGATION DELAY, RECEIVER DELAY, AND RELATIVE MOTION OF SATELLITE

As the orbital elements of satellites are not decoded by the TRANET I system receiver, propagation delays from satellites to the receiving antenna are calculated by using the data of Cartesian coordinates of the satellite 30190 which are supplied by the Topographic Center of DMA. Fig. 2 shows an example of propagation delays in vacuum space which were calculated by the precise ephemeris of the satellite 30190.

Receiver delays are measured by transmitting a simulated signal into a space near the receiving antenna at few days interval. During the period before November 28, 1976, the AF type of tracking receiver was used and adjustments of IF circuit were made so as to give a constant delay in the receiver. Thereafter the IF phase-lock tracking receiver was used and only measurements of receiver delay were made without frequent adjustments. Delay time of the IF phase-lock tracking receiver is shown in Fig. 3.

Delays of time signals emitted from satellites can be calculated according to the Lorentz transformation. But classical treatment is sufficient, for the radial velocity of satellites (v) relative to the station fixed on the Earth is very small against the light velocity (c). Then time delay caused by the motion of satellite relative to the tracking station is estimated as

$$(v/c) \cdot (\text{propagation time}) \lesssim 0.5\ \mu s$$

Time delay of this kind is corrected, although this is small.

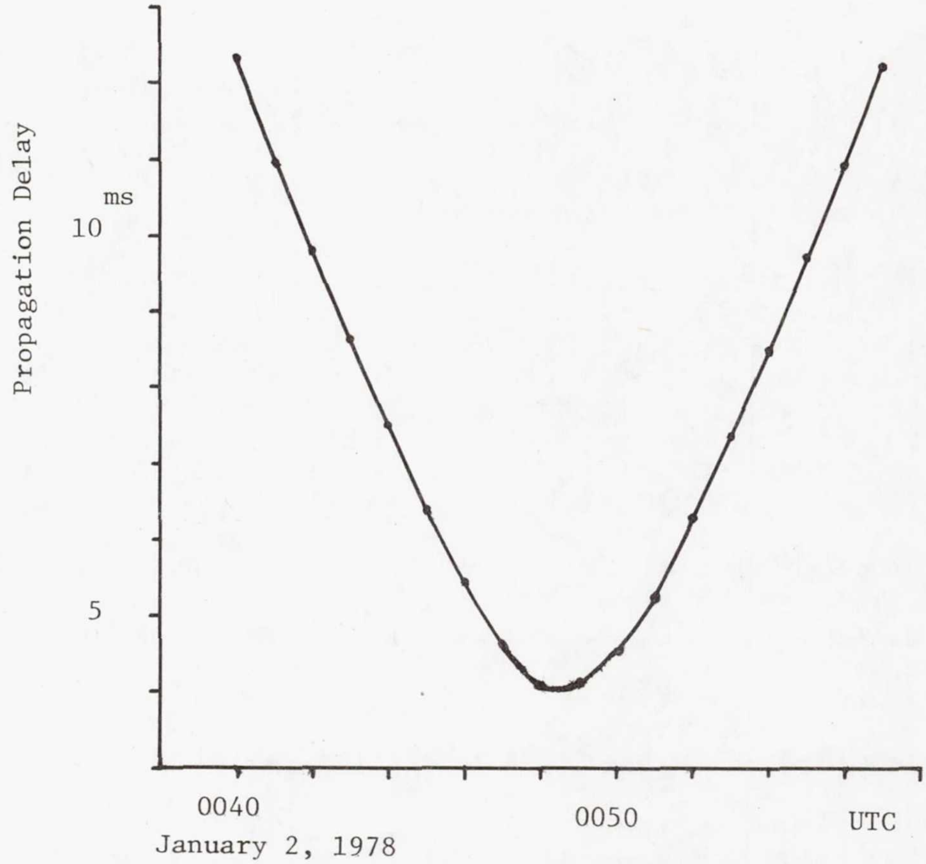


Fig. 2-An example of propagation delays in vacuum space which were calculated from the precise ephemeris of the satellite 30190.

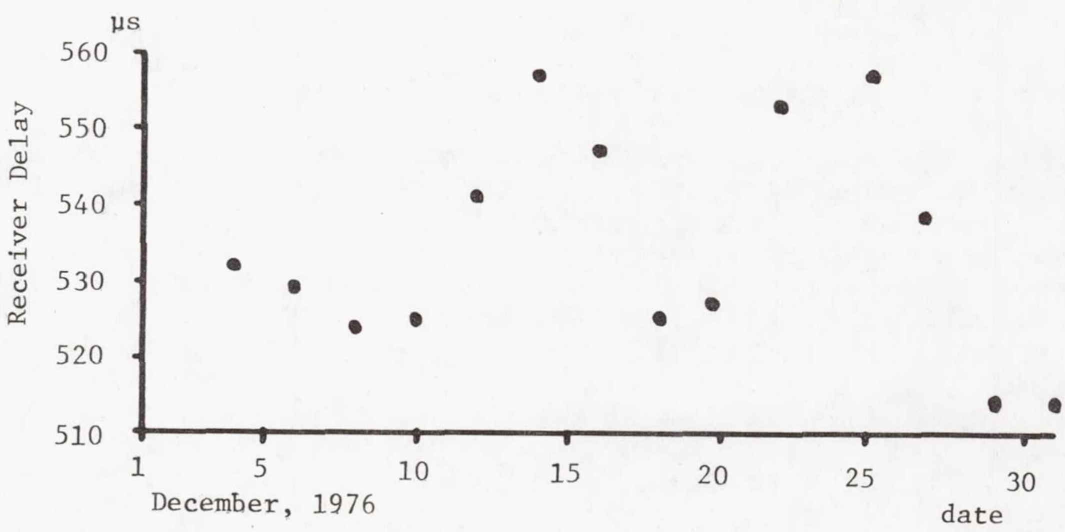


Fig. 3-Receiver delays of the TRANET I IF phase-lock tracking receiver.

RESULTS

Time differences between the satellite 30190 and the station clock are shown in Fig. 4, where the corrections for propagation delay, receiver delay, and motion of satellite mentioned above are made.

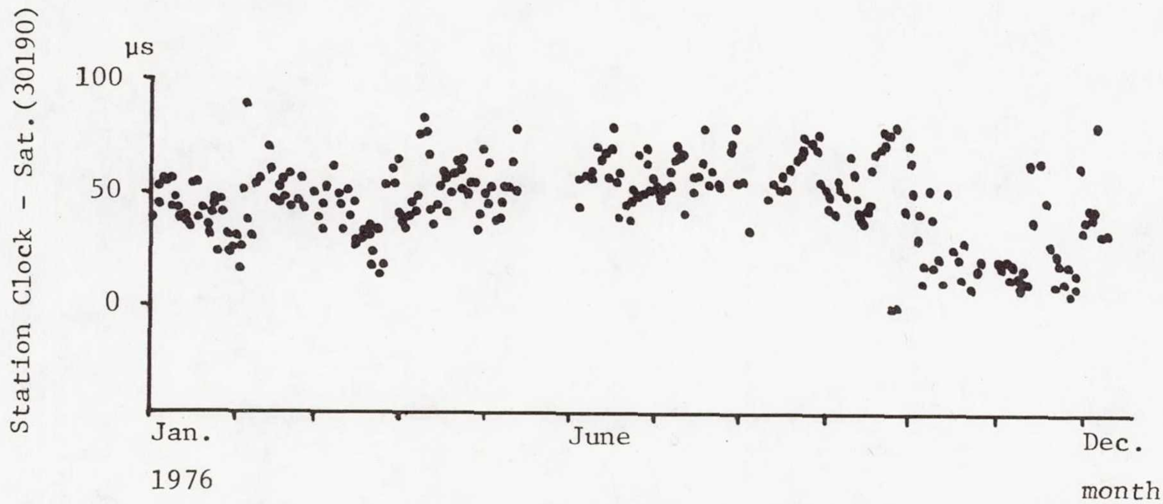


Fig. 4-Emission time of the satellite 30190 observed by the station clock (rubidium atomic clock).

In this calculation, receiver delays are corrected by three constant values in the three periods, respectively, as follows;

775702 µs before 23 January 1976,
775763 µs until 28 November 1976,
775966 µs after 28 November 1976,

where system delay of 775434 microseconds and 406 Hz circuit delay are included.

The raw data included some extremely deviated values, and these data were rejected by a fixed range filter to pass only the data which were in the range from 0 to 350 microseconds. The refined data were proved to have the standard deviations of ±19 microseconds. The above measurements were made with the station clock of which deviations were about ±1.5 microseconds.

In order to estimate the intrinsic error of the clock comparison via satellite, we tried to remove all the

errors of the station clock and satellite-borne clock (see Fig. 5) and receiver delay only for the period after November 28, 1976.

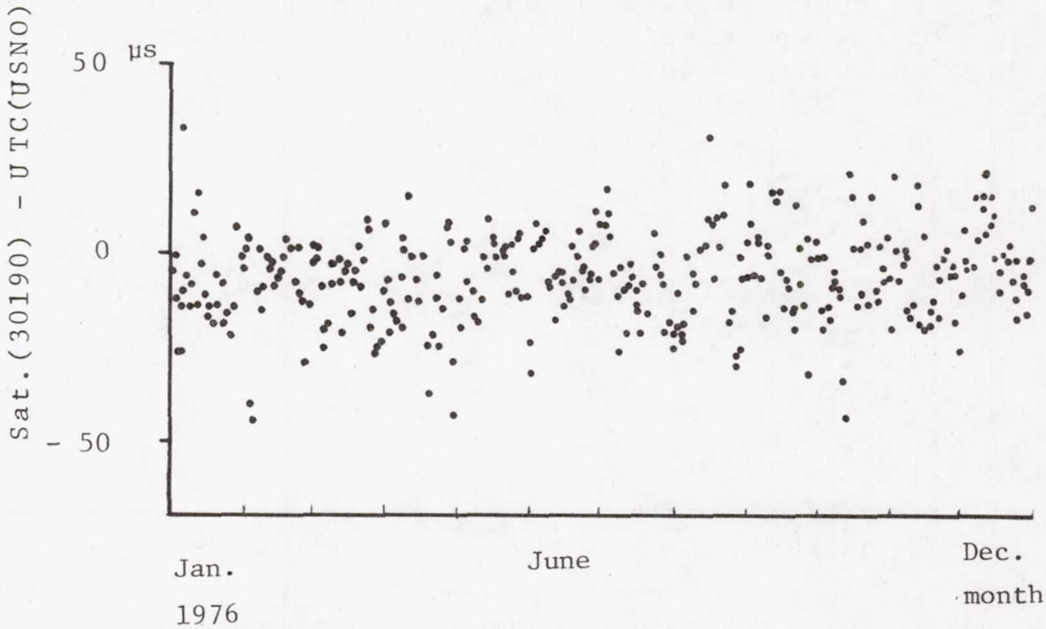


Fig. 5-Time differences between satellite-borne clock (30190) and UTC(USNO). This graph was plotted from the data in Transit Satellite Report, Series 17 which was published by the U.S.Naval Observatory.

The final form of clock comparison was reduced to UTC(ILOM) -UTC(USNO) (see Fig. 6). These two time scales are maintained by the cesium clocks and fluctuations in UTC(ILOM)-UTC(USNO) can be ascribed to timing error aroused by the satellite timing system. Results are summarized below with standard deviations in microseconds;

Station Clock - Sat.(30190)	±19
Sat.(30190) - UTC(USNO)	±12
UTC(ILOM) - Station Clock	± 1.5
Receiver Delay	±14
UTC(ILOM) - UTC(USNO)	±15.5

Standard deviations were reduced from ±19μs to ±15.5μs, showing that only slight improvements were attained. This may due to instability of the receiver delay and/or to inappropriate correction of satellite clock.

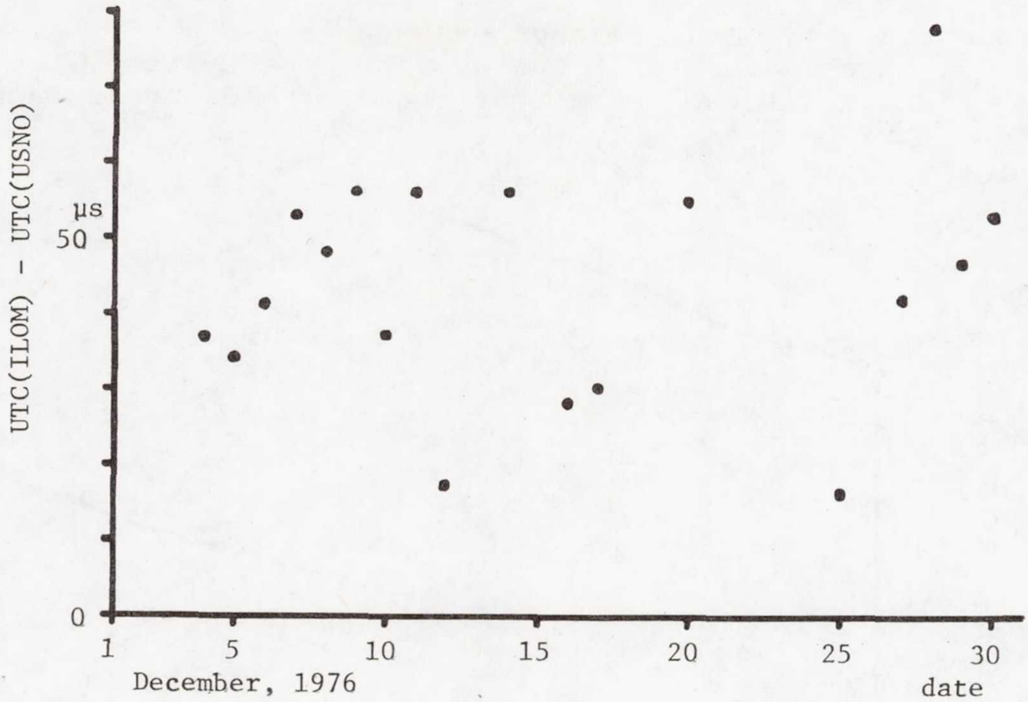


Fig. 6-UTC(ILOM) - UTC(USNO) via Transit satellite 30190.

TIME SYNCHRONIZATION VIA LORAN-C SIGNAL

The Northwest Pacific Loran-C chain is available in the vicinity of Japan. Most institutes in Japan have been receiving the master station Iwo-jima with the standard deviation of less than ±0.1μs. Receiver delay can be measured with sufficient accuracies, but propagation time seems to be hard to estimate with high accuracies. Propagation time from Iwo-jima to the monitoring station of the chain (Fuchu) was once determined by USNO as 4070.0μs from a calculation combined with a transportation experiment with an atomic clock. RRL when it was in Midori-cho, Koganei-shi, Japan had calculated the propagation time based on above value as 4122.5μs. On the other hand, time synchronization between RRL and ILOM has been made with the aid of a portable clock as well as Loran-C receptions. The difference of propagation times between these two stations and the Iwo-jima station was determined as 1192.4±0.24μs by five clock transportation experiments. Then, propagation time from Iwo-jima to ILOM was obtained as 5314.9μs.

Thus, the quantity of UTC(ILOM) - UTC(USNO) is derived from the Loran-C receptions and the data "Daily Phase Values,

Series 4" which is published by USNO (see Fig. 7). By comparing these values with the one obtained by satellite timing signals, it was found that there is a discrepancy of about 35μs.

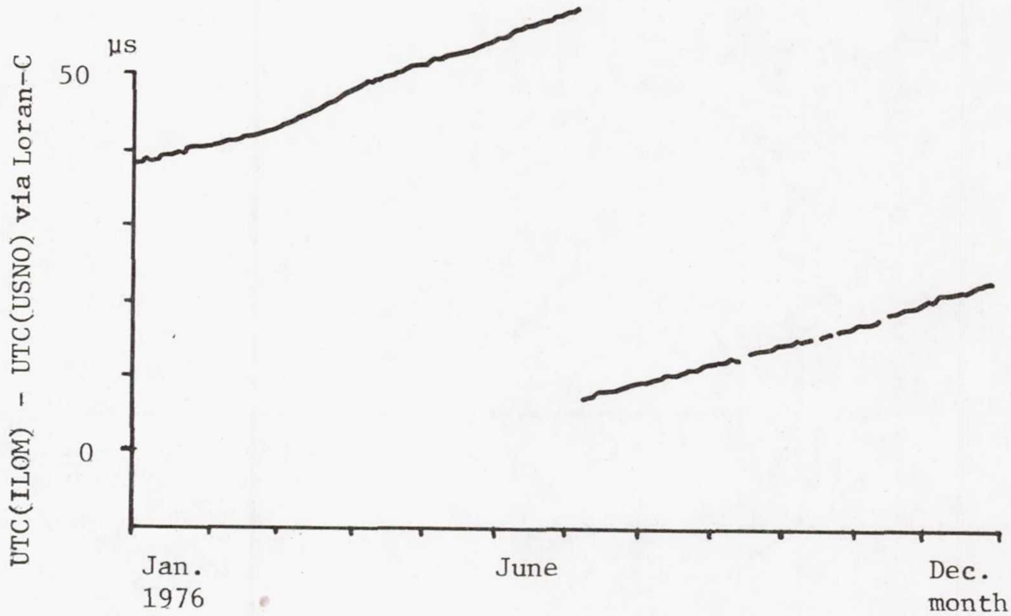


Fig. 7-UTC(ILOM) - UTC(USNO) via Loran-C.

DISCUSSION

The ionospheric effect on radio wave propagation was not corrected in the above results. The order of the effects on propagation delay will be estimated here. As the receiving frequency is 400MHz, geomagnetic field and collision of electrons with neutral gas can well be neglected. Then the optical length (τ) is calculated according to the formula $\tau = \int n ds$, where the integration must be done along the propagation path and n is the refractive index which is in relation with the plasma frequency (f_p) and the operating frequency (f) as $n^2 = 1-(f_p/f)^2$. The plasma frequency is related with electron density (N) as $f_p{}^2 = 80.6N$ in MKS unit system. Then the optical length can be estimated from the equation $\tau = \int(1-80.6N/f^2/2)ds$ by using a model ionosphere (Tsuchiya 1976) for the N(h) profile. A numerical calculation was made for the satellite which is on observer's zenith, yielding 0.1μs and 0.05μs in daytime and nighttime respectively. The distance to the satellite which is on the horizon will be four times as large as the one

when a satellite is on the zenith. So, maximum propagation delay may amount to 0.4μs and 0.2μs, respectively.

On the other hand, the mean value of station clock - satellite was obtained as 48.95 ±1.39μs and 48.70 ±1.77μs for the daytime and nighttime period, respectively. That is, reception error is above the ionospheric effect, so we can find no significant differences in ionospheric effect between daytime propagation and nighttime one from the above results. Nevertheless we may safely say that the ionospheric effect produces no errors larger than one microsecond when the solar activity is moderate.

At TRANET stations only the data which are obtained when the satellite is near to the closest approach (C.A.) are used for time synchronization purpose. All the data that fall in the range from 0μs to 350μs were used, in our case. Dependency of delay time of timing signals upon the doppler shift of satellite were examined for each datum point in whole passes obtained in 1976 (see Fig. 8).

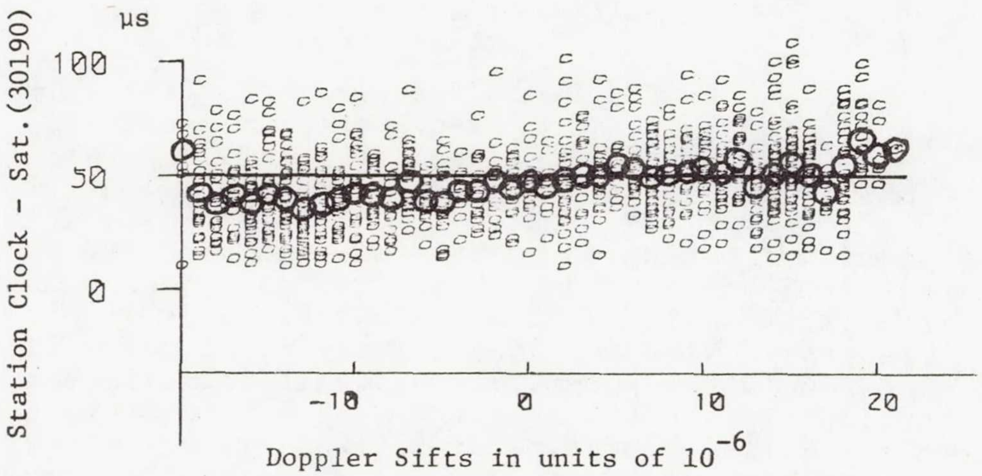


Fig. 8-Station clock - Satellite 30190 which was observed every two minutes. The abscissa is taken as Doppler shift in units of 10^{-6}. Average value is shown by large circles.

Scattering of data are relatively small near C.A., and furthermore there is a tendency that received signals advance by about 10μs for pre-C.A. period and vice versa for post-C.A. period. This tendency does not differ distinctly whether the reception was made during daytime or nighttime

periods. The physical explanation of this tendency is left unsolved even if we take into consideration the tropospheric refraction effects, since these are the order of one microsecond at most.

CONCLUDING REMARKS

A timing experiment via the Navy Navigation Satellite for the year of 1976 was shown. Our time comparison has shown that fluctuations of the obtained data have the standard deviation of $\pm 16\ \mu s$. This is almost the same order as the reported values by Hunt and Cashion (1978) and Cashion et al. (1978). But there is a discrepancy of $35\ \mu s$ as compared with the data obtained by the Loran-C reception. There might be some problems in delay time measurement. Furthermore, fluctuations in timing pulses may be pretty large, since the band width of the receiver is narrow.

ACKNOWLEDGEMENTS

Our thanks are due to Mr.M.Aihara and Dr.I.Okamoto for their discussion and assistance in improving expressions.

REFERENCES

[1] BLAIR,B.E., Time and frequency dissemination: An overview of principles and techniques, NBS Monograph 140, 233-313, 1974.

[2] CASHION, R.E., G.A. HUNT, and J.D. LAVANCEAU, Precise time recovery from Transit satellites, paper presented at Symposium on Time and Frequency XIXth General Assembly of URSI, 1978.

[3] FREQUENCY STANDARD SECTION and KASHIMA BRANCH, Report on clock pulse synchronization experiment via Relay II satellite, J.Radio Res. Labs., 12, 63, 311-316, 1965.

[4] HUNT, G.A. and R.E. CASHION, A Transit satellite timing receiver, Proc. Ninth Annual Precise Time and Time Interval Planning Meeting, 153-166, 1978.

[5] TSUCHIYA, A., Geodetic application feasibility of new technics in position astronomy, J. Geod. Soc. Japan, 22, 4, 264-274, 1976.

[6] U.S.NAVAL OBSERVATORY, Daily phase values and time differences, Series 4 and Transit satellite report, Series 17.

[7] YAMAMOTO, M., K. HARADA, and Y. SABURI, A time comparison experiment performed by SSPA via ATS-1, J. Radio Res. Labs., 23, 10, 85-103, 1976.

QUESTIONS AND ANSWERS

MR. LAUREN RUEGER, Johns Hopkins University, Applied Physics Lab:

Before I open this paper for comments from the audience, I would like to make a couple of comments myself. The first is that the small shift he saw in this last curve is characteristic of what we observed in the tracking loop characteristics of the VCO. He was using a fairly early model Nikode-type receiver that has a fairly simple transfer function for the tracking loop. The lags in that would give him the 10 microseconds I think he is observing.

The second comment is that during 1977, following this data, we did an experiment in making time transfers between the U.S. Naval Observatory and the National Bureau of Standards in which we had very carefully calibrated the receiver delay, to a resolution of 10 nanoseconds. And in using that, we discovered, buried in the data that we now provide through Bulletin 17, a possibility of a 50 microsecond bias because of the uncertainty of the receiver delay.

If you take the 50 microseconds from this source and the 35 microsecond discrepancy this man found, they are in the same direction and compensating. His data is really within his experimental error. We should tell him someday.

Page intentionally left blank

DOWN TO EARTH RELATIVITY

Irwin I. Shapiro
Department of Earth and Planetary Sciences
and
Department of Physics
Massachusetts Institute of Technology
Cambridge, MA

ABSTRACT

The basic concepts of the special and general theories of relativity are described. Simple examples are given to illustrate the effect of relativity on measurements of time and frequency in the near-earth environment.

Down to Earth Relativity

I. Introduction

As almost everybody knows, next March 14th is the 100th Anniversary of the birth of Albert Einstein, and so many celebrations and symposia are planned for 1979 that I fear all will become violently ill from an overdose of relativity well before mid year. For now, I would like to distill some of the salient aspects of both the special and the general theories of relativity and to relate them to clocks and frequency standards. After describing the basic concepts of special and general relativity, I'll discuss the size of the relativistic effects near the earth and the level of their experimental verification to indicate how well one might be able to rely on general relativity.

II. Special Relativity

Special relativity is partly concerned with the perceptions of observers viewing rods and clocks in uniform motion relative to one another (and not accelerating with respect to some "absolute" inertial frame which we won't worry about here). A key idea in Einstein's development of this theory involves the concept of simultaneity. If, as Newton assumed, there was a universal time coordinate that applied throughout all space, then there is no problem in our agreeing on a definition of the simultaneity of two events: We simply compare the readings of our "universal" clock. If the readings are the same at each event place, we agree that those events took place simultaneously.

If there is a spatial coincidence between two points, then there's again no problem agreeing on a definition of simultaneity because the points are co-located. We can use the same watch at the same place to see whether the events occur at the same time. That's no problem, with or without a universal time.

If there were spatial separation between two events, and if we could communicate between those two separate spatial points with infinite speed, then again, we'd all agree there would be no problem in deciding whether or not the events were simultaneous.

However, if we have spatial separation and the communication speed is limited by the speed of light, as Einstein thought, then, there is a problem. The definition of simultaneity is no longer intuitively obvious. In fact, as a simple, down to earth, example can show, even with a rea-

sonable definition, there is not necessarily agreement on simultaneity among observers moving relative to one another.

Now, let us define simultaneity for events at spatially separated points with communication between them possible via light signals. Concentrate for a moment on some given frame (Figure 1).

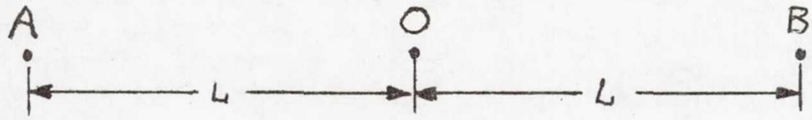


Figure 1

Any Single Frame

We're concerned about whether or not events that occur at points A and B in this frame are simultaneous. We can go to the midpoint between these two points, a distance L from A and B, and we can say that when an observer at point A sees an event he or she (hereafter "he" for economy) immediately transmits a light signal toward O and when the observer at point B sees an event he also immediately transmits a light signal to O. If these two light signals arrive at O simultaneously, then we say that the events A and B occurred simultaneously. That's a reasonable definition of simultaneity.

Now suppose we have two frames in relative motion. Consider, in particular, a down to earth example: the ground and a train (Figure 2).

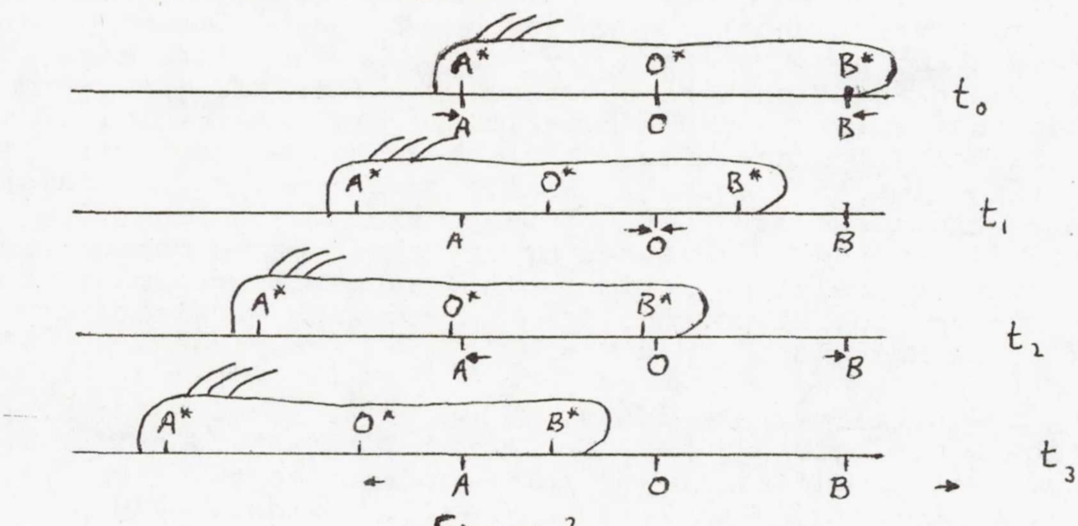


Figure 2

Two Frames In Relative Motion

(Any who know my flying habits, know that I have a special spot in my heart for trains even though they do tend to run a bit late.) Let us single out two points, A and B, on the ground, and set an observer halfway between. We follow the same procedure on the train so that at a certain instant, say t_o, we have our train point A* coincident with A, our observer O* coincident with O, and B* coincident with B. Suppose that an event occurs at both A and B at t_o as measured in the "ground" frame and that light signals are transmitted from A toward O, and from B toward O at that instant. As the light signals travel, the train is, of course, moving, say in the direction from B to A. Thus, the light signal from A is going to arrive at O* while the light signal from B is still travelling toward O*. A little while later, the two signals arrive at O, simultaneously, so the observer at O would say events A and B occurred simultaneously. But our observer at O* would not agree because he received the signals from A* and B* at different times. So one may conclude that if an event is simultaneous as measured according to our definition in one frame, the event will not necessarily be simultaneous as measured in another frame. Of course, there is nothing special about any one frame: the events could as well have been arranged to appear simultaneous to the observer at O* as to the observer at O.

In fact, one may conclude more generally that simultaneity is dependent on the motion of the observer. Einstein also thought that all observers moving uniformly with respect to one another should be equally valid observers so there should be no preferred (inertial) frame. Further, he felt that there should be no region of space-time singled out as more important than any other; he therefore assumed that space and time were homogeneous. This assumption implies that a linear transformation determines the relation of (Cartesian) space-time coordinates in one frame with respect to those in another. Einstein made one more assumption: the speed of light, c, is constant, such that the same value would be measured by any observer no matter what his state of relative motion. This seemingly provocative assumption had, of course, been upheld with exquisite accuracy in Michelson's and Morley's 1887 experiment.

To quantify these ideas, Einstein utilized a transformation which actually had been derived somewhat earlier, although with a different and inferior intellectual foundation, by Lorentz - the Lorentz Transformation:

S*   S

$$
\left.
\begin{aligned}
x^* &= \frac{x - vt}{[1 - (v/c)^2]^{1/2}} \\[2ex]
t^* &= \frac{t - (v/c^2)\, x}{[1 - (v/c)^2]^{1/2}}
\end{aligned}
\right\} \qquad (1)
$$

This transformation relates the coordinates x and t, defined in a frame S, to the corresponding coordinates in a frame S*, where S* has a velocity v with respect to S. One can see that these transformations are linear in x and t and in x^* and t^*. Since neither S nor S* is "preferred", we should be able to invert the equations and obtain the same description, and, indeed, we do, except for the sign inversion of the velocity:

S   S*

$$
\left.
\begin{aligned}
x &= \frac{x^* + vt^*}{[1 - (v/c)^2]^{1/2}} \\[2ex]
t &= \frac{t^* + (v/c^2)\, x^*}{[1 - (v/c)^2]^{1/2}}
\end{aligned}
\right\} \qquad (2)
$$

Because S* has a velocity of v with respect to S, S has a velocity minus v with respect to S*. There is nothing in the transformation that singles out any one frame as special, despite the theory's being called "special" relativity (for a different reason, discussed below).

Let us now turn to the question of clocks. What in particular, does special relativity say about clock rates? If there is a clock at rest in the frame S* and one measures, from frame S, the interval between two "ticks", the result will be different from the corresponding measurement made by an observer in frame S*. The numerical relation between these measurements made in different frames is given by

$$\Delta t = \frac{\Delta t^*}{\sqrt{1-(v^2/c^2)}} \quad . \tag{3}$$

In other words, the observer in S thinks the interval between ticks is longer than does the observer in S*. The ratio of the two intervals is given by the Lorentz factor:

$$\sqrt{1 - (v^2/c^2)} \quad . \tag{4}$$

Similarly, since these are symmetric situations, if the clock had been in S, and the observer in S* were to measure the interval between two ticks, a similar relation would be obtained:

$$\Delta t^* = \frac{\Delta t}{\sqrt{1-(v^2/c^2)}} \quad . \tag{5}$$

The point is that, with the clock in S*, one is actually comparing it to a series of (identical) clocks distributed in S which the clock in S* passes as it moves with respect to S; similarly, with the clock in S: it moves with respect to the fixed (identical) clocks in S*. There is no paradox in the relationship being symmetric. One may conclude from this analysis that a clock always appears to run fastest in its own frame. When an observer is at rest with respect to the clock, he thinks the clock is running faster than when he is in uniform motion with respect to that clock. This effect has been verified very well in the measurement of the lifetime of unstable elementary particles. Such particles in cosmic rays, and in accelerators, often move with velocities _v_ very close to _c_, and this lifetime enhancement factor can then be very large because the Lorentz factor tends to zero and its inverse to infinity. Studies of mu-mesons have verified this effect with very high accuracy.

III. General Relativity

The theory of relativity we just discussed is special in the sense of being restricted. It is silent on the subject of gravitation; it is concerned primarily with physics in (inertial) frames moving uniformly with respect to one another. Einstein felt that concern was not sufficient; he

wanted to introduce gravitation. Einstein was unhappy with Newton's theory of gravitation which had existed unchallenged for about two centuries. Newton's theory did, of course, have one small problem. There was a minute, but annoying, discrepancy between the observations and the theory which became noticeable in the late 1850's and was quite well established by the early 1900's. Einstein was not upset about Newton's theory only because of a mere disagreement with observations; his concern was a matter of principle.

Einstein did not accept Newton's theory because it implied action at a distance. In this theory, the force felt by body A due to body B depended on the location of body B at the very instant that body A felt the force. But if no signal can travel faster than the speed of light, how is body A to know where body B was located at that instant? This aspect was a severe drawback to Newton's theory in Einstein's mind and he set about the development of an alternative. The process took about a decade. The main principle upon which he based this general theory of relativity is the so-called "principle of equivalence".

One can state this principle in various ways. A usual way is to state that the effect of a gravitational field locally is indistinguishable from an inertial acceleration. The example usually given is that of an "Einstein elevator". Suppose a laboratory is enclosed in an opaque small elevator and placed in a gravitational field, such as on the surface of the earth. The scientists inside feel the force of gravity but cannot unequivocally identify it as such. They may do any physics experiments and obtain numerical results. However, suppose now the laboratory were taken away from the earth and accelerated uniformly with a rocket. If the scientists in the laboratory were to repeat all their experiments, the principle of equivalence states that they will get exactly the same numerical answers, provided that the inertial acceleration is exactly equal to the gravitational acceleration.

Another statement of the principle of equivalence can be given in terms of the ratio of gravitational to inertial mass. Gravitational mass is the mass that appears on the right hand side of the equation that expresses Newton's law of gravity:

$$F = \frac{G m_g M_g}{r^2} \tag{6}$$

where G is Newton's constant of gravitation, m_g is the (gravitational) mass of the body being acted upon, M_g is the (gravitational) mass of the body attracting m_g, and $\underline{r}$ is the distance between them. The inertial mass is the coefficient of the acceleration, $\underline{a}$, in Newton's law of motion:

$$F = m_i a \tag{7}$$

The ratio of these two masses, according to the principle, is independent of the composition of the bodies and independent of the mass of the bodies. It is a universal constant. This principle, although not so named, was also accepted by Newton. In fact, he was the first to verify it quantitatively, achieving an accuracy of approximately 1 part in 1,000.

What can we infer from this principle of equivalence? One of the things Einstein inferred was that the trajectory of a particle could depend only on the geometry of space and time. By the principle, the trajectory did not depend on the particle itself, on its composition, or on its mass (except for the "back reaction" which I ignore here). It doesn't matter whether we have a pea, a flashlight, or whatever; it will move on the same trajectory because it will be affected in the same way as any other mass. Thus, Einstein reasoned that one could talk about the trajectories being merely a property of the geometry, and having nothing to do with the particular object that was moving along the path.

What determines the geometry? Einstein felt that the geometry should be determined solely by the mass, or, more precisely, the mass-energy, distribution in the universe. But isn't it contradictory to say that the path of the particle doesn't depend on the particle, only on the geometry, and that the geometry depends only on the mass distribution? Certainly the particle is part of the mass distribution. Yes, but if one considers the particle to have an infinitesimally small mass, it won't affect the geometry, and to that extent, these statements are consistent. But this "closed loop" aspect is a key to Einstein's theory.

Einstein may have been guided in developing his "field theory" for gravitation by analogy with Newtonian physics. In Newtonian physics, one obtains the gravitational potential from the mass distribution. In other words, the gravitational potential everywhere in space is determined by the mass distribution. In fact, the potential, Φ, is determined by Poisson's Equation:

$$\nabla^2 \Phi = 4\pi\rho \quad , \tag{8}$$

where ρ is the mass density. Only the gravitational potential appears on the left side and only the mass (density) on the right side. This equation is a linear, second-order, partial differential equation for Φ. Einstein in effect generalized this purely spatial expression, to an analogous space-time expression that also allowed the geometry to be non-Euclidean. He used Riemannian geometry and developed an analogous equation where, on the right side, the mass density is replaced by the energy-momentum tensor:

$$G_{\mu\nu}(g_{\mu\nu}) = 8\pi T_{\mu\nu}; \ \mu,\nu = 1 \to 4. \tag{9}$$

As in the Newtonian case, only the right side contains the "mass" terms; only the left side contains the "geometry" dependence. The geometry here is defined in a metric space. The so-called "metric tensor" $g_{\mu\nu}$ in essence expresses the "connection" between neighboring points in this space-time:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{10}$$

The interval ds is the "distance" between two infinitesimally separated points in the space-time. To evaluate ds^2, one sums over all values of the two indices μ and ν which run from 1 to 4 and correspond to the three spatial and the one temporal dimension. In Cartesian coordinates, in Euclidean three-dimensional space, $ds^2 = dx^2 + dy^2 + dz^2$; Equation (10) is the generalization for a Riemannian metric space.

Einstein made other assumptions, namely that this energy-momentum tensor is a conserved quantity in a sense analogous to the conservation of energy and momentum in Newtonian physics. Further, he limited the derivatives on the left side to second-order derivatives of $g_{\mu\nu}$, in analogy to the second-order derivatives on the left side of the Newtonian equation (8). With those assumptions, one can uniquely determine the left side up to a term proportional to the metric tensor. The coefficient of this term, the so-called cosmological constant, Einstein first took to be zero, a position he deviated from later when he thought the universe was static; still later, he greatly regretted this temporary deviation. (It is now generally assumed that the cosmological constant is zero.)

Because of symmetry ($T_{\mu\nu} = T_{\nu\mu}$), Equation (9) represents only 10 independent equations, not 16. These are Einstein's field equations which he used as the basis for calculations. In Newtonian Physics, the field equations were not enough. Equation (8) indicates how the gravitational potential can be determined, but it doesn't tell one how to calculate the paths of light rays and particles. In fact, Newton never said anything, as far as I know, about the effect of gravity on light rays. As for the effect of gravity on massive objects, Newton had a separate assumption, his equally well-known law of motion, given in Equation (7). In relativity, the corresponding equations are the equations for geodesics in four-dimensional space-time. A very intriguing aspect of the general relativistic formulation is that a separate assumption for the equations of motion does not seem to be needed; the equations of motion follow from the field equations themselves. The basic reason that makes this result possible, though by no means guaranteed, is that the field equations of general relativity are non-linear. The Newtonian field equation, by contrast, is linear. The terms hidden in $G_{\mu\nu}$ in Equation (9) are, in fact, non-linear expressions in terms of the metric tensor $g_{\mu\nu}$.

IV. Magnitude of Relativistic Effects

What of the magnitude of the relativistic effects we might expect? We know, as Einstein also knew, that Newtonian physics is a very good approximation, at least in our neighborhood. So the Newtonian equations must be, in some sense, the first approximation for the solution to the relativistic equations. Deviations from Newtonian physics appear in terms proportional to v^2/c^2 as we saw from the Lorentz Transformation; in the general theory of relativity, deviations appear in terms proportional to the factor, GM/c^2r. The quantity GM/c^2 has the dimensions of length and is often denoted by r_o and called the gravitational radius of the body. We can evaluate r_o near the sun, say, to determine the order of magnitude of the relativistic effects there that are due to gravitation. We find that, for the sun, $r_o \simeq 1.5$ km; by contrast, the radius of the sun is about 700,000 km. Thus, we can expect relativistic effects to appear at the level of two parts per million.

What about effects near the earth, which are of more direct concern for us? We find that the gravitational radius of the earth is near half a centimeter. In other words, the earth would have to be compressed down to half a centimeter before it would turn into a black hole. The radius of the earth is about 6 x 10^8 centimeters, so relativistic

effects near the surface of the earth could be expected to be on the order of eight parts per billion, not terribly large.

Let us now try to describe the relativistic effects quantitatively. To solve the field equations to determine the metric tensor $g_{\mu\nu}$, is no easy job. There are very few problems that have been formulated where $g_{\mu\nu}$ can be determined in closed form. The most famous one, solved by Schwarzschild very shortly after Einstein published his theory, is for a spherically-symmetric, static mass distribution. The solution exterior to that mass can be written as:

$$ds^2 = -\left(1 - 2\alpha\frac{GM}{rc^2} + 2\beta\left(\frac{GM}{rc^2}\right)^2 + \ldots\right)c^2dt^2$$

$$+\left(1 + 2\gamma\frac{GM}{rc^2} + \ldots\right)(dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2), \quad (11)$$

where, as stated above, ds^2 is the infinitesimal space-time interval and where the non-vanishing components of the metric tensor are the coefficients of dt^2, dr^2, etc. These coefficients, as here, are often written as a power series in GM/c^2r. In general relativity, the parameters, α, β, and γ, don't appear; they are identically one. The higher-order terms, indicated by "...", do not appear either; they are identically zero in general relativity. The reason for writing the metric in this "generalized" form is to facilitate the testing of the theory. By a least-squares match of the predictions of the theory to the results of observations made, say, in the solar system, one can estimate the values of these parameters. If the estimates turn out to be unity to within experimental uncertainty, we conclude that the observations are consistent with general relativity. If they aren't, then general relativity is in trouble.

Given the metric tensor and the equations of motion, one can then calculate explicitly the paths of test particles and light signals. The calculations are a bit intricate; one cannot in general obtain "closed-form" solutions. One often uses a perturbation expansion in powers of r_o/r where the first approximation represents the Newtonian solution and the next higher approximation, the so-called post-Newtonian solution.

V. Simple Examples of Relativistic Effects on Frequency and Time

Let us, finally, turn to the predicted relativistic effects on the frequency of light signals and on clocks. We will treat first a very simple example to show how one can use elementary reasoning to obtain an answer, without employing the full armamentarium of general relativity. We'll need only to apply the principle of equivalence. Thus, suppose we have a transmitter and a receiver that are stationary but separated. Let the receiver, or observer, be on the surface of the earth; let the transmitter, at an altitude H above the observer, transmit a signal with frequency $\underline{f}$ (see Figure 3). The question is, "What frequency does the observer measure?"

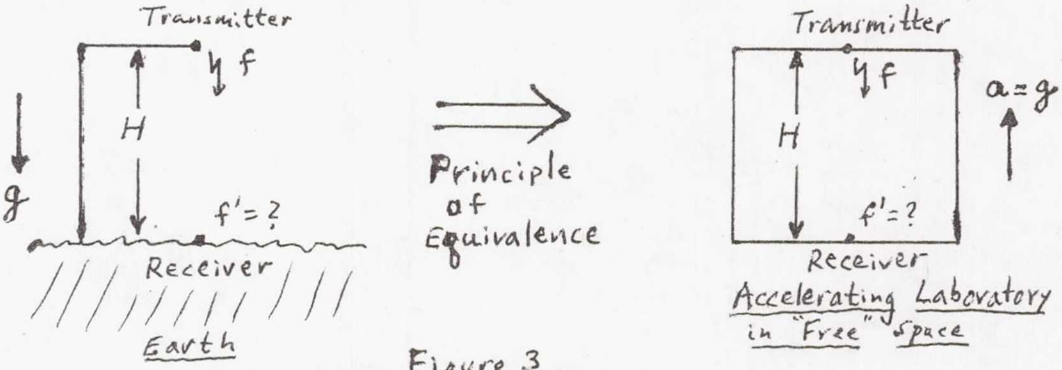


Figure 3
Application of Principle of Equivalence

An easy way to answer this question is to use the principle of equivalence. The system, or laboratory, we set up is equivalent to another where we replace gravity by an acceleration: We accelerate the laboratory at a value $\underline{a}$, equal in magnitude to the acceleration $\underline{g}$ of the earth's gravity. We keep the observer and transmitter separated by the same distance H. At some instant, the transmitter sends a signal which the observer receives a short time, Δt, later. Let the velocity of the observer, at the instant of reception of the signal, relative to his and the transmitter's velocity, at the instant of transmission of the signal, be Δv. The value of Δv will be equal to the acceleration of the laboratory multiplied by the time interval between transmission and reception. Thus, using the principle of equivalence,

$$\Delta v = a\Delta t = g\left(\frac{H}{c}\right) \tag{12}$$

where Δt is just the time taken by light to travel the distance H (Δt = H/c), and where a = g. These are all approximate relations, valid to the first order in the small quantities. The frequency shift, $\Delta f \equiv f' - f$, of the observed frequency relative to the transmitted frequency, f, is like a first-order Doppler shift and is given by

$$\frac{\Delta f}{f} \simeq \frac{\Delta v}{c} = \frac{gH}{c^2} \tag{13}$$

where we have substituted from Equation (12). This change in frequency represents, in fact, a violet shift.

Thus, the transmitter, at altitude H, sends a signal at frequency f and the observer receives a signal with a frequency greater by Δf. We note that the change in gravitational potential between transmitter and receiver is just the change in $-GM_{\oplus}/R_{\oplus}$, the gravitational potential for the earth:

$$\Delta\Phi = \Delta\left(-\frac{GM_{\oplus}}{R_{\oplus}}\right) \simeq \left(\frac{GM_{\oplus}}{R_{\oplus}^2}\right)\Delta R_{\oplus} \simeq gH \tag{14}$$

where $M_{\oplus}$ and $R_{\oplus}$ are the mass and radius, respectively, of the earth; and $\Delta R_{\oplus}$ is equal to H. The fractional change, $\Delta f/f$, in frequency and the accumulated difference, $\Delta\tau$, in apparent clock readings after elapsed time τ are given by:

$$\left.\begin{aligned} \frac{\Delta f}{f} &\simeq \frac{\Delta\Phi}{c^2} \\ \Delta\tau &\simeq -\frac{\Delta\Phi}{c^2}\,\tau \end{aligned}\right\} \tag{15}$$

In other words, if the observer had a clock identical in construction to that governing the transmitter, and if the observer knew the value of the transmitted frequency, as determined at the transmitter, by the clock there, the observer would infer that his clock was losing time relative to the clock in the lower gravitational potential of the transmitter. Of course, this "relativistic" loss can easily be taken into account in any comparison.

Let us consider another example. Suppose a frequency standard were in a circular orbit about the earth, and sup-

pose, incorrectly, that the first-order Doppler shift and the earth's rotation were negligible. Suppose, further, that a signal of frequency, f_s, is transmitted by the satellite and received on earth, and that the frequency of that signal is measured on earth, with equipment governed by a clock identical in construction to the clock in orbit that governed the transmission of the signal. Under our assumptions, the frequency, f_e, measured on earth will be related to f_s by:

$$f_e = \frac{[1 + (2\Phi_e/c^2)]^{1/2}}{[1 + (2\Phi_s/c^2) - (v_s/c)^2]^{1/2}} f_s \quad , \tag{16}$$

where the subscripts $\underline{s}$ and $\underline{e}$ refer to conditions at the satellite and on the earth, respectively. Thus, the difference, $f_e - f_s$, in frequency is determined by the motions of the bodies and by their gravitational potentials. Recall from Equation (2) that for the motion itself, we have the factor $(1 - v^2/c^2)^{+1/2}$; but here, where we are considering frequency rather than time, this factor enters with the plus half power rather than with the minus half power. As we saw in the first example, although not in this more exact form, the gravitational potential also affects the frequency; the effect was linear in (Φ_s/c^2) with a coefficient of unity. This result can be recovered here, for $(\Phi_s/c^2) \ll 1$, by expansion of $(1 + 2\Phi_s)^{1/2}$. Since (Φ_s/c^2), (Φ_e/c^2), and (v_s^2/c^2) are all small near the earth, we expand the right side of Equation (16), rearrange, and obtain (with the aid of conservation of energy):

$$\frac{\Delta f}{f} \equiv \frac{f_e - f_s}{f_s} \simeq \frac{3GM_\oplus}{2(R_\oplus + H)} - \frac{GM_\oplus}{R_\oplus}$$

$$\simeq 3.5 \times 10^{-10} \left[\frac{3R_\oplus}{(R_\oplus + H)} - 2\right] \quad , \tag{17}$$

where H is the altitude of the satellite. The ratio $\Delta f/f$, the apparent fractional change in frequency measured by the observer on the surface of the earth is thus of the order of a few parts in 10^{10}, where for H less than half the radius of the earth we observe a violet shift, and for H greater than half the radius of the earth we observe a red shift. Above half an earth radius, the effect of the motion dominates over the effect of the gravitational potential, and vice versa, below half an earth radius. With the combination of the motion and the gravitational potential

effects, we would measure either a violet shift or a red shift, depending simply on the altitude of the satellite. Were we to observe from a lower potential, that is, from a position higher above the earth than the satellite, we would measure a red shift. Remember, however, that this entire development must really be modified for the observer's motion and for the first-order Doppler shift, both of which were ignored in this example.

VI. Validity of the General Theory of Relativity

Now let us address briefly the question of whether or not general relativity is a valid theory. It is clear in principle that at some level general relativity must "break down", because it is incompatible with quantum mechanics. No one has yet been able to formulate a satisfactory quantum theory of gravity, although there are some good ideas currently being explored. As one makes observations on a more microscopic scale, quantum mechanics plays an increasingly important role. At what length scale will quantum gravity actually be important? One answer is based on the evaluation of the "fundamental" length that can be formed from the gravitational constant, the speed of light, and Planck's constant, h, which is a measure of the importance of quantum phenomena. This length is called the Planck length and is given by:

$$L_P \simeq \left(\frac{\hbar G}{c^3}\right)^{1/2} \simeq 1.6 \times 10^{-33}\text{cm} \quad , \tag{18}$$

where, in accordance with convention, the "slash" on h denotes division by 2π. It is clear for present PTTI purposes that one need not worry about such length scales. It will be a long time before anyone will conceive of practical experimental procedures that will expose what happens at these length scales. Quantum theories of gravity currently under study center on so-called "super gravity", which tries to unite general relativity and quantum mechanics in a "higher level" theory for which general relativity will be the appropriate macroscopic limit. Testing the validity of these ideas is hopelessly beyond present experimental capabilities.

In the macroscopic world of the solar system, relativistic effects are very small. In addition, they have been verified by measurements to one percent or better. The relativistic effects of motion and gravity on clock rates, in particular, have been verified to approximately one hundredth of one percent already. A relativistic effect on trajectories, the prediciton of a non-Newtonian advance of

the perihelion position of Mercury's orbit, has been verified to about half a percent. The predicted deflection of light rays, and the predicted increase in echo delays, have been verified to the order of one percent, and a few tenths of one percent, respectively.

There is no problem, in principle, in applying the general theory of relativity to the solar system, and, in particular, to the earth environment at a useful level of accuracy. The situation is all very well defined by the principles of the theory. Unfortunately, how to apply these principles is not always so clear to those who try. As one consequence, apparent paradoxes have appeared in the literature, as well as many other errors. But, at the level of accuracy of interest to PTTI, these are the problems of those doing the calculations, and not the problems of the theory. The theory is quite reliable and often useful at this level of accuracy.

QUESTIONS AND ANSWERS

DR. CARROLL ALLEY, University of Maryland:

I think it is appropriate for this audience to realize that the first practical applications of Einstein's ideas in actual engineering situations are with us in the fact that clocks are now so stable that one must take these small effects into account in a variety of systems that are now undergoing development or are actually in use in comparing time worldwide.

It is no longer a matter of scientific interest and scientific application, but it has moved into the realm of engineering necessity. So talks like this are very important to try to acquaint the community with these fundamental principles, because the uncertainties have, indeed, arisen in lack of understanding of what is going on, rather than in the basic ideas.

DR. SHAPIRO:

Yes, in fact I left out one slide where I meant to show what the accumulated effect, say, in a day would be if you took two identical clocks, put one on the ground and one in the spacecraft in orbit around the earth at some nominal altitude.

Of course, we can cancel it out as we saw, but what would be the order of magnitude of the accumulated difference in the readings of the two clocks per day? And it is about 20 microseconds. So it can be quite substantial.

Of course, that is a little bit of a spoof since we don't yet have such extremely stable absolute standards, so if you put a clock in orbit and just measure its rate in orbit, then you would, in effect, automatically correct for these relativistic effects, provided it was a circular orbit and provided certain other things were true.

But when one gets down to the tens of nanoseconds level, and one worries about eccentric orbits and various other things, then it is true that these effects, small as they are, are not negligible compared to the accuracy that you can achieve with clocks.

The first really practical application that I know of that people are worried about is in the GPS system, where the effects are of the order of tens of nanoseconds for some of the applications.

DR. ALLEY:

For the GPS, albeit a 12 hour orbit, it is 38 microseconds per day.

DR. SHAPIRO:

That is true. But I say that you can get rid of that very easily by the redefinition of rate.

DR. ALLEY:

Yes.

DR. SHAPIRO:

But you still have to worry even in comparisons within a day of the order of tens of nanoseconds.

DR. ALLEY:

If I may be permitted one more comment: In the summer of '77 we actually carried out the Einstein falling-elevator experiments using the earth falling towards the sun. We transported clocks essentially from the floor to the ceiling by carrying them from the northern hemisphere to the southern hemisphere at the time of the summer solstice, when the axis is tilted toward the sun. We verified for clock rates that the potential of the sun does not effect the clock rates between floor and ceiling in the freely-falling elevator earth. Thank you.

DR. SHAPIRO:

There are many experiments, as I alluded to, that verified various aspects of general relativity. I felt I couldn't do justice to all of them, and therefore I did justice to none of them.

DR. CHARLES MARTIN, Defense Mapping Agency:

I would like to make one comment here because I think it's quite important in terms of our potential utilization of the global positioning system. I don't think there is any question about the microsecond errors if you do not take them into account.

But I think it is certainly important that we all realize that the capability, the theory, is adequate to take into account relativity errors to the level of, say, 20 or 30 nanoseconds.

DR. SHAPIRO:

No. To much better than that. My main message was the theory makes very specific predictions and they have been verified to a small fraction of one percent as far as clocks are concerned.

So, simply on the experimental verification level, you can believe them to the sub-nanosecond level. But as far as the theory is concerned, there is no good reason to believe it breaks down there just because you haven't tested below there. There is no theoretical reason that it should break down just below. And it does make very specific predictions. The problems arise, as I said, when people don't fully understand the theory when they try to use it in their calculations.

MR. ALLAN, National Bureau of Standards:

I again think for this audience, along the lines Professor Alley mentioned, that for the GPS user in the future, because the earth is

spinning, these effects become very significant. If you synchronize two clocks on the surface of the earth via portable clock and via satellite (by GPS), and ignore that the earth is spinning, assuming the Einstein synchronization technique, you can make errors of the order of hundreds of nanoseconds. So one has to be careful.

DR. SHAPIRO:

That is right. One has to be careful. But I am saying that the theory is very clear. I could work out any example, including the spinning earth, including flying clocks westward against the direction of earth (as was done already) and eastward with the direction. And there are differences there, because you are adding to or subtracting from the velocity of rotation of the earth. All of these things have been worried about and have been calculated and there is no problem, as long as you really understand the theory that you are applying.

MR. THOMAS MCCASKILL, Naval Research Laboratory:

We have a talk this afternoon in which we will present some results with the NTS satellites. In view of the high amount of interest that has been shown on the relativistic effects, we will bring a couple of slides that Mr. Buisson presented last year, which show the difference in frequency between a cesium clock measured on the ground and a cesium clock that was placed in orbit, which verified the first order relativistic effect.

DR. ALFRED KAHAN, Rome Air Development Center:

In your opinion, then, is there any experiment that still needs to be done to further prove the general theory of relativity with satellites, flying clocks? Or is the theory so good that we have confirmed to the one-percent or half-percent level that we don't need any more experiments?

DR. SHAPIRO:

I am a firm believer that physics is an experimental science and when one has the opportunity to test to a higher level of accuracy one should, provided it doesn't cost a major fraction of the gross national product.

And one has to draw some reasonable position there between doable but hugely expensive and do-able but not such a great gain. I believe in experiments if you can make an order of magnitude gain in the experimental limit: It is worth a reasonable amount of money.

If you are going to make a ten percent gain, I personally wouldn't bother doing the experiment. There are some effects of general relativity that haven't been observed at all at any level that are important.

For example: The dragging of inertial frames due to the spinning of the massive body were predictions worked out from

general relativity as long ago as 1918. They have never been verified because the effects are very small.

There are several possible ways of getting an experimental handle on this with earth experiments, including flying spinning gyroscopes and so forth, but they are technically very difficult and very expensive to perform, and it is not clear yet that we are really ready to do that.

DR. ALLEY:

I would like to adopt a slightly different stance. The confusion in the understanding of the fundamental principles is widespread even among authorities.

I mean, there are recently published papers in the literature making predictions coming from people who should know better. For example, on this falling earth experiment I mentioned, one of the leading theorists in Europe in general relativity published in <u>Physics Letters</u> the flat statement that clocks would run at different rates at the North Pole and South Pole at the time of the solstices.

This is flat wrong, which he now admits. But there is a tremendous amount of intuition that is lacking in understanding general relativity, which we have in electricity and magnetism. And I would submit that the performance of clock experiments that we are now able to do will contribute vastly to developing this kind of intuition.

In a certain sense the clocks in gravitational fields are analogous to magnetic filings in magnetic fields. And it is quite important to do these experiments when one is able to do them.

DR. SHAPIRO:

I don't like to disagree with my colleague, but I find that I must disagree strongly with what Professor Alley just said. I find that no amount of experiment can really take people away from wrong notions. For example, the twin paradox has created fanatics in great numbers and no amount of experiments quells that at all.

As far as theoretical physicists like the one to whom Professor Alley alluded, and whom he didn't mention and whom I won't mention, he was perfectly well convinced that he had made an error simply on a theoretical basis. It didn't take an experiment to convince him that he made an error.

It was perfectly clear that he just didn't apply properly the relativistic principles. Many people, if they are reasonable, can be convinced by the theoretical arguments, and having exposed their wrong step, they admit it.

The non-fanatics will be convinced by the theory, and the fanatics won't be convinced by anything.

MODERN METHODS FOR THE DETERMINATION OF POLAR MOTION AND UT1

William E. Carter
National Geodetic Survey
National Ocean Survey, NOAA
Rockville, Md. 20852

ABSTRACT

This paper is organized into two major divisions according to the topics: polar motion and UT1. Each division is introduced with a brief review to provide a minimal perspective for readers unfamiliar with the subject area. The applications of Doppler satellite observations, laser ranging to artificial satellites and the Moon, and astronomic radio interferometry to monitoring polar motion and UT1 are discussed. Emphasis is placed on detailing how and what each method is capable of measuring, fundamental limitations are noted, and the present status of the development of each method is reviewed.

The paper concludes with a summary of the author's evaluations of the various methods as candidates for the next generation international polar motion and UT1 monitoring service.

INTRODUCTION

The "classical" methods of monitoring polar motion and UT1 have been based on visual, photographic, and photoelectric observations of optical stars. The temporal and spatial resolutions and accuracies of these methods have been limited by such factors as; an inability to fully correct the observations for the effects of the Earth's atmosphere, inaccuracies in the relative positions and proper motions of the stars, the limited number and poor distribution of observatories, and instrumental imperfections. Further refinements of the classical methods, some of which involve the application of modern technology, are continuing and are expected to yield significant improvements. However, profoundly different methods, which have developed as outgrowths of space exploration activities, promise an order of magnitude improvement in our ability to monitor polar motion and UT1. It is these space-age methods, i.e., Doppler satellite observations, laser ranging to artificial satellites and the Moon, and astronomical radio interferometry that are

discussed in this paper.

POLAR MOTION

Review

Polar motion is the motion of the Earth's instantaneous pole (axis of rotation) with respect to a reference point fixed to the crust of the Earth.

The theoretical basis for the existence of polar motion was presented by Euler in 1765, but the motion was not detected observationally until the late 1800's. S. C. Chandler discovered that the observed motion was actually the result of two primary components: a revolution of the true pole around the principal moment of inertia axis counterclockwise when viewed from the north, with a period of 1.2 years; and an annual revolution, also counterclockwise (Chandler, 1891). The 1.2 year period of the first component (now commonly referred to as Chandlerian motion) did not agree well with the much shorter period predicted by Euler's work. The discrepancy was quickly explained by S. Newcomb as being due to the elasticity of the Earth (Newcomb, 1891). The annual term is produced by the continuous redistribution of mass in meteorological and geophysical processes.

The motion of the pole is not totally predictable from a simple two-component model. Unexpected changes in the magnitude and direction of the motion occur, that result in a requirement to monitor the motion on a continuing basis.

Regular monitoring of polar motion was undertaken by the International Latitude Service (ILS) in 1899, and has continued without interruption until today. The ILS system uses the differential zenith distance method (Hoskinson and Duerksen, 1947) of determining latitude with visual zenith telescopes (VZT). The stations are all located very near the same parallel of latitude (39° 08' N) so that the same star pairs can be observed from all observatories. The mean pole position defined by the ILS observatories for the period 1900-1905 has been adopted as the Conventional International Origin (CIO).

In 1962 the ILS was reorganized, according to resolutions of the International Astronomical Union, and the International Polar Motion Service (IPMS) was founded (Yumi, 1964). The IPMS continues to publish polar positions based only on the ILS observatories, but it also publishes values derived from a combination of VZT, Photographic Zenith Tube (PZT), astrolabe, and transit circle observations from approximately 75 observatories.

In 1955, the Rapid Latitude Service (RLS) was established, by action of the IAU, under the direction of the Bureau International de l'Heure (BIH), to predict the coordinates of the pole and provide time corrections with very short delays. The individuality of the RLS has since been abandoned and the rapid service is now provided as a routine function of the BIH.

In 1968, the BIH adjusted the positions of their contributing observatories, predominately the same observatories that are included in the IPMS system, to insure the coincidence of the BIH pole with the CIO. Since that time, the BIH reference system has been maintained independently from the ILS system. In 1972, the BIH began to include pole position information obtained by Doppler satellite observations in their solutions. The Doppler values used are the two-day solutions of the Transit navigational system observations, presently published by the Defense Mapping Agency. The methods used to combine the Doppler data with the optical data are detailed in the <u>BIH Annual Report for 1976</u>.

Pole positions derived from Doppler observations of artificial satellites are available from as early as 1967, but the earliest data are of lower quality than the post 1972 data. The developmental work of the Doppler Polar Motion Service (DPMS) was accomplished at the Naval Weapons Laboratory (Anderle, 1973). The Defense Mapping Agency (DMA) took over operational responsibility in April 1975 (Oesterwinter, 1978). The Doppler polar positions are available directly from DMA, and are also published in <u>U. S. Naval Observatory Time Service Publication Series 7</u>.

To briefly summarize, polar motion values are determined and distributed today by the IPMS, BIH and DMA. The IPMS utilizes only optical data, the BIH utilizes a combination of optical and Doppler satellite data, and DMA utilizes only Doppler data. Monthly means are usually quoted as having uncertainties in the 20 to 40 cm range, but the positions published by the different services often differ by 1 to 2 meters.

Many questions still remain unanswered even after almost 80 years of continually monitoring polar motion. Some of these questions cannot be answered unless significant improvements are made in the spatial and temporal resolutions of the observations. An improved monitoring system, based on more modern methods, is badly needed. Candidate methods are: Doppler satellite observations, laser ranging to the Moon and artificial satellites, and astronomic radio interferometry.

Polar Motion Determinations by Doppler Satellite Observations

The material presented in this section has been extracted primarily from papers by Anderle (1973) and Oesterwinter (1978).

Radio signals suitable for Doppler observations are transmitted by U. S. Navy Navigation System satellites. The satellites are in nearly circular polar orbits at heights of about 1,000 km. They continuously transmit at two carrier frequencies, 399.968 MHz and 149.988 MHz (nominal values). The oscillators typically drift a few parts in 10^{11} per day. Both frequencies are generated from the same oscillator to facilitate the determination of ionospheric refraction effects.

Pole positions are obtained as part of the bi-daily updating of the orbit of each satellite. The gravity field model and the positions of the base stations are held fixed in a least squares solution which estimates the x and y coordinates of the pole, six constants of orbital integration, one drag scaling factor, a frequency and tropospheric scaling factor for each satellite pass, and the coordinates of any new points being surveyed. The bi-daily solutions from as many as five different satellites are combined to derive 5-day mean positions of the pole.

The Doppler pole positions are determined in the "Doppler network" coordinate system. In 1970, an attempt was made to make the origin of the Doppler coordinate system close to the CIO by estimating the coordinates of the base stations in a solution in which the gravity field coefficients and the BIH pole positions were held fixed. The network does vary with time due to station failures, modifications and upgrades, and the augmentation of the 17 to 20 base stations by one to ten, or more, temporary stations during various operational campaigns.

The standard error of the pole positions vary considerably depending upon the distribution and number of observations combined in each solution. Oesterwinter (1978) concludes that the standard deviation of a two-day polar coordinate solution is now better than 40 cm, and for a five-day mean, under 20 cm. The dominant source of error is believed to be residual errors in the gravity field model.

Polar Motion Determinations by Satellite Laser Ranging

The material presented in this section has been extracted primarily from papers by Kolenkiewicz et al. (1977) and Smith et al. (1978).

Many artificial Earth satellites have been equipped with retroreflectors to facilitate tracking by laser ranging systems. Polar motion can be determined from satellite laser ranging from a single station, if an accurate satellite ephemeris is available, or a network of stations.

In the case when only a single tracking station is in operation, only one component of polar motion, i.e., the component along the station meridian, can be monitored. The procedure is to establish a precise

reference orbit by tracking the satellite for a reasonable period of time, say a month or so, and then compare subsequent observations made over periods of perhaps 6 to 12 hours, to this reference orbit. Of particular interest is the apparent change in inclination of the orbit, since changes in the latitude of the tracking station are reflected as apparent changes in that parameter. Of course, in order to extract the changes in latitude, any real changes in the inclination of the satellite must be taken into account.

The strongest determination of the inclination of the orbit is obtained when the tracking station is located near the northern or southern apex of the orbit. The satellite is then moving along an east to west (or west to east) track to the north (or south) of the tracking station when it is observed.

Initial experiments to determine polar motion by satellite laser tracking were conducted by the NASA Goddard Space Flight Center during a 5-month period in 1970. The experiment used ranges to the Beacon Explorer C satellite. The pole positions showed residuals, with respect to BIH values, having an rms deviation of about 1 meter. A very fundamental difficulty with the single station method just described is the requirement for a reference orbit that remains very precise over periods of months to years.

If a "network" of ground stations is operated, for which a consistent set of coordinates is known, the requirement for the reference orbit is eliminated and both the x and y components of the pole position can be determined in the "network frame of reference." This multistation case corresponds to the Doppler methods previously described.

Smith et al. (1978) presented the first results of the network approach using the Laser Geodynamics Satellite (LAGEOS) which was launched on May 4, 1976. The LAGEOS satellite was specifically designed to have a very stable orbit. The satellite is in a high orbit (12,265 km) and has a high mass to surface area ratio, which greatly reduces the effects of such perturbing forces as solar radiation, Earth albedo, air drag, and the high frequency components of the Earth's gravitational field.

Smith and Kolenkiewicz analyzed LAGEOS data collected by a network of seven stations during the period of May through December 1976 and derived 5-day mean values of x and y with most of the estimated standard deviations ranging from 0.01 to 0.02 arcsecond ($\sim$30 to 60 cm). These results are available in tabular form (Smith, 1978).

Polar Motion Determinations by Lunar Laser Ranging

The material presented in this section has been extracted primarily from Harris and Williams (1977).

An error in the latitude of a lunar laser ranging observatory results in a range error which is relatively constant near zero hour angle, but does depend on the lunar declination, approximately as the sine of the difference between the latitude of the observatory and the lunar declination.

$$\Delta r = r\Delta\phi \sin(\phi-\delta) \tag{1}$$

where r is the range from the observatory to the lunar reflector; δ is the declination of the lunar reflector; ϕ is the geocentric latitude of the observatory; $\Delta\phi$ is the change in the geocentric latitude of the observatory due to a change in the position of the pole; and Δr is the change in range.

When the Moon's declination is nearly equal to the latitude of the observatory, the error in latitude contributes little to the range error, but as the Moon's declination moves away from the observatory's latitude, the error in latitude does contribute to the range error. The signature has a period equal to the lunar cycle. Harris and Williams analyzed the McDonald lunar laser ranging data to determine if there was a reasonable chance of extracting daily polar motion values from that single observatory data set. They concluded that there was not. Polar motion values based on lunar laser ranging apparently will have to await multiobservatory operations.

Polar Motion Determination by Radio Interferometry

In radio interferometry the range-difference and the time rate of change of the range-difference from the radio sources to two (or more) observatories are determined. When extra-galactic sources, such as quasars, are observed the sources are at such great distances that the radio wave fronts arriving at the interferometer are essentially planar. Any purely translational motions of the interferometer or rotations about an axis parallel to the interferometric baseline are not detectable. For a single baseline, it is possible to determine two of the three angles necessary to express the orientation of the Earth in the frame of reference defined by the radio sources. To determine all three angles, two nonparallel baselines are required.

The sensitivity of radio interferometry to variations in the orientation of the Earth can be estimated from the following equations:

$$\Delta X = -(\Delta\Theta)\, Y - x\, Z \tag{2}$$

$$\Delta Y = (\Delta\Theta)\, X + y\, Z$$

$$\Delta Z = x\, X - y\, Y$$

x and y are the components of the pole position, in radians, relative to the CIO; $\Delta\Theta$ represents (UT1-UTC), also in radians; X, Y, Z are the Earth fixed coordinates of the baseline; and ΔX, ΔY, ΔZ are the changes in the baseline vector components caused by x, y and $\Delta\Theta$.

If the baseline has a substantial Z component, changes in the position of the pole will cause significant changes in the equatorial components of the baseline (X,Y) which will be reflected in the sinusoidal signatures of the delay and delay rates.

If the baseline is nearly parallel to the equatorial plane, i.e., $Z \simeq 0$, sensitivity to polar motion comes solely from the ΔZ term. A small variation in Z causes a change in the delay that varies with the declinations of the sources. It is therefore quite feasible to determine one component of polar motion from an equatorial baseline. In fact, excellent determinations of the x component of polar motion have been obtained from the nearly equatorial Haystack-Owens Valley radio interferometer. These results are particularly noteworthy because they agree closely with Doppler derived values, the rms of the differences being $\simeq$30 cm, and display obvious systematic trends relative to the BIH values (Robertson et al. 1978).

There are two approaches to implementing astronomic radio interferometry; connected element interferometry (CEI), and very long baseline interferometry (VLBI). The underlying principles are not different, but the technological methods used to bring the signals detected at the two telescopes together for processing cause significant differences in the operational characteristics of the interferometers and in the dominant error sources.

CEI baselines are presently limited in length to a few tens of kilometers by the ability to maintain the phase stability of the connecting data link. VLBI baselines are limited in length to a few thousand kilometers by the size and shape of the Earth. The inherent angular resolution of an interferometer is directly proportional to the baseline

length - i.e., the longer the baseline the better the angular resolution.

However, the theoretical resolution of the interferometer is not presently the limiting constraint on the accuracy to which the polar motion can be determined. Since polar motion is an angular measurement, any instabilities of the two telescopes forming the interferometer degrade the determination as the ratio of the radius of the Earth to the length of the baseline. For baselines of a few kilometers to a few tens of kilometers, for which CEI is presently feasible, this multiplicative factor is of magnitude 10^2 to 10^3, and antenna distortions due to gravitational and wind loading and temperature variations, and local crustal deformations become critical. Atmospheric "seeing" also appears to be a serious problem (Hargrave and Shaw, 1978). For the much longer baselines used in VLBI these factors decrease in significance and local oscillator instabilities are presently the limiting constraint. As higher performance oscillators are developed, the ultimate limiting constraints are likely to become atmospheric effects, possible variations in the structure of the radio sources and tectonic motions.

In October 1978, the U. S. Naval Observatory (USNO) initiated a program to use the 37 kilometer CEI located at the National Radio Astronomy Observatory (NRAO) at Greenbank, West Virginia. NRAO personnel perform the observations specified by USNO, and the data reduction and dissemination are done by Washington based USNO and Naval Research Laboratory (NRL) personnel. The USNO/NRL group estimates that the NRAO interferometer may eventually be capable of determining polar motion to 0.01 arcsecond (30) cm over an averaging period of a day or so.

The National Ocean Survey of the National Oceanic and Atmospheric Administration has begun a project to establish and operate a three-station network of permanent observatories to monitor polar motion (and UT1) by VLBI. The project designation is POLARIS (POLar-motion Analysis by Radio Interferometric Surveying). Project POLARIS is described in some detail in Carter et al. (1978). Computer simulations indicate that the POLARIS system will be capable of determining the x and y components of polar motion to better than 10 cm over an averaging period of 8 hours.

UT1

Review

The Bureau International de l'Heure (BIH) has monitored the rotation of the Earth since 1912, utilizing observations from a large number of stations (presently about 80) distributed around the world. The BIH is presently the only international service which provides UT1 data. Independently determined values are also published in the USNO Time Service Series 6.

The Earth does not rotate at a constant rate, but exhibits periodic, secular, and irregular variations. The primary periodic terms have annual, semi-annual, 27.55 day and 13.66 day periods. Sudden variations in the length of day of several milliseconds over a period of a few days have been observed (Smylie and Mansinha, 1968).

Just as with polar motion, the present methods of monitoring UT1 do not have sufficient angular or temporal resolutions to satisfy modern scientific requirements. In the following sections, candidate methods for improved UT1 determinations are examined.

UT1 Determinations by Artificial Satellite Observations

In order for the rotational orientation of the Earth to be determined in an inertial frame of reference by artificial satellite observations, it is necessary that any perturbations of the satellites' orbital plane be predictable over the time span of interest. Over time spans of several months to years, the uncertainties for even the most stable satellites, such as LAGEOS, will grow to unacceptable levels. For this reason it is widely agreed that artificial satellite observations by Doppler, laser ranging, VLBI or any other method are not suitable for long term Earth rotation studies. However, satellite observations could be used to monitor short-term variations in the Earth rotation which could then be combined with observations of a different type having the desired long-term stability.

Smith et al. (1978) have investigated the use of LAGEOS laser ranging observations for the determination of UT1. The limiting perturbations appear to be Earth albedo and ocean tides. They estimate that by the early 1980's the modeling capabilities will likely be such that it will be possible to derive UT1 with uncertainties of a few tenths of millisecond over periods of three months. Silverberg (1978) has suggested that the pairing of LAGEOS and lunar laser ranging data would be a reasonable marriage of convenience. The LAGEOS data would provide unbroken short-term coverage, while the lunar data would provide the

long-term frame of reference. It would be desirable if the participating observatories could do both types of range measurements.

Another combination that has been suggested is VLBI observations of artificial satellites, such as the NAVSTAR constellation, and extragalactic sources. The satellite signals could be made of sufficient strength that relatively simple and inexpensive ground stations and data processing systems could be used. The more expensive observations of the extragalactic sources would be minimized, and yet still provide the long-term stable frame of reference.

UT1 Determinations by Lunar Laser Ranging

A constant error in the assumed longitude of a lunar laser ranging station will produce a range residual, with respect to the lunar ephemeris, which varies as the sine of the lunar hour angle.

$$\Delta r \approx \Delta \mathrm{UT0} \cos \delta \sin H \qquad (3)$$

Δr is the change in range from the observatory to lunar reflector due to an error in longitude (ΔUT0); δ is the declination of the lunar reflector, H is the hour angle of the lunar reflector.

If lunar laser ranging data are available over a large enough span of hour angle to allow the sinusoidal signature to be well estimated, UT0 can be determined for the observatory. Using polar motion information from another source, such as the IPMS or BIH, or observations from two or more properly located lunar laser ranging observatories, it is possible to obtain UT1. Just as in the case of artificial satellites, the lunar ephemeris must not drift, relative to an inertial frame of reference, if the UT1 determinations are to remain accurate over long periods of time.

A significant number of lunar laser range measurements have to date been made from only one observatory - the University of Texas McDonald Observatory. These data have been analyzed by several investigators to extract Earth rotation information (Stolz et al. 1976; Harris and Williams 1977; King et al. 1978).

King et al. compared UT1 determinations by lunar laser ranging to smoothed BIH values, to which fortnightly and monthly corrections had been added to account for tidal effects largely removed by the BIH

smoothing procedure. After removal of the mean differences, the rms of the remaining differences was 2.1 msec. It should be noted that the UT1 values derived from the lunar laser ranging data by different investigators, differ significantly. For example, the King et al. vs. Stolz et al. UT1 values have an rms difference of 0.8 msec. The sources of these differences are being investigated.

UT1 Determinations by Radio Interferometry

Elsmore (1973) points out that an equatorial interferometer ($Z \approx 0$) can measure UT1 directly - uncorrupted, to first-order, by polar motion. Equations 2 make this quite apparent. For $Z \approx 0$, the terms containing x and y (the components of the pole position) vanish from the equations for ΔX and ΔY. For this reason, it would be desirable in designing a UT1 monitoring network to utilize equatorial baselines.

Elsmore (1973) reported the results of UT1 determinations with the 5 km equatorial CEI at Cambridge between August 1972 and May 1973. The rms scatter relative to BIH values was approximately 6 msec. The averaging time for each determination was generally 12 hours.

USNO estimates that it will be able to determine UTO using the NRAO CEI with an accuracy of 1 msec on a daily basis during 1979, and hopes to substantially improve that accuracy over the next few years.

Robertson et al. (1978) reported the results of 14 VLBI experiments conducted between September 1976 and May 1978 using the Haystack-Owens Valley interferometer, in which UT1 values were determined. The rms difference with respect to the BIH values (with added fortnightly and monthly terms) was 1.6 msec, after removal of the mean difference. Comparative studies of the VLBI and lunar laser ranging UT1 values are in progress.

NOS computer simulations for the POLARIS network indicate that it will be capable of determining UT1 to ± 0.1 millisecond in an averaging period of eight hours or less. The experimental results of the Haystack-Owens Valley interferometer, reported by Robertson et al. (1978), add credibility to the simulations.

SUMMARY

The only truly operational usage of any of the modern techniques for the determination of polar motion or UT1 has been the Doppler satellite determinations of polar motion. Even in this case, the polar motion information has been a by-product of a program having other primary goals, and very little effort has been made to optimize the network configuration, observing schedules, or instrumentation for determining polar

motion. The Doppler satellite method suffers the disadvantages that it relies on the availability of functional satellites, which have limited lifetimes and must be replaced from time to time, and that it would not be an adequate method for the long-term determination of UT1. The author believes that this limits the Doppler satellite method to a transitional role that will prove very useful in verifying the initial results of methods better suited to long-term usage.

Viewed independently, laser ranging to artificial satellites and the Moon, each have serious deficiencies. Even using very special satellites such as LAGEOS, it appears that satellite methods will only be able to provide measurements of UT1 over periods of a few months before errors due to rotation of the orbital plane become excessive. Lunar laser ranging suffers from technological complexities (only one observatory has been made to operate reliably after the first full decade of experiments) and difficulties in obtaining measurements within a few days before or after a new Moon. The combined usage of satellite and lunar laser ranging data, as suggested by Silverberg, certainly offers some promise but the cost of operating enough stations to ensure reliability during periods of poor weather may still disqualify these methods.

Astronomic radio interferometry has several very desirable attributes: a three-station network (the minimum number of stations required to form two baselines) can determine both components of polar motion and UT1; spatial resolutions of 10 cm can be achieved in averaging periods of eight hours or less; observations can be made during inclement weather; the radio sources form the most nearly inertial frame of reference presently known; the radio sources have unlimited lifetimes, and are equally available to all users. The choice between CEI and VLBI techniques depends very heavily on the maximum length of the baseline that can be operated in the CEI mode. For baselines of less than several hundred kilometers the task of cleansing the CEI data of purely local effects appears, in the author's opinion, to be insurmountable.

Of the presently available methods, VLBI observations of extragalactic radio sources appear to be the best choice for use in the next generation international polar motion and UT1 service.

REFERENCES

Anderle, R. J., 1973: "Determination of Polar Motion from Satellite Observations," Geophysical Surveys 1, pp. 147-161, D. Reidel Publishing Company.

Bureau International de l'Heure 1977, Annual Report for 1976.

Carter, W. E., Robertson, D. S., Abell, M. D., 1978: "An Improved Polar Motion and Earth Rotation Monitoring Service Using Radio Interferometry," Proceedings of International Astronomical Union Symposium No. 82, Time and the Earth's Rotation, Cadiz, Spain, May 1978.

Chandler, S. C., 1891: "On the Variation of Latitude," The Astronomical Journal No. 249, pp. 66-70.

Elsmore, B., 1973: "New Method for Direct Determination of UT1," Nature Vol. 244, pp. 423-424.

Hargrave, P. J., Shaw, L. J., 1977: "Large-Scale Tropospheric Irregularities and their Effect on Radio Astronomical Seeing," Monthly Notices Royal Astronomical Society 182, pp. 233-239.

Harris, A. W., Williams, J. G., 1977: "Earth Rotation Study Using Lunar Laser Ranging Data," in Scientific Applications of Lunar Laser Ranging, Edited by J. D. Mulholland, pp. 179-190, D. Riedel, Kingman, Mass.

Hoskinson, A. J., Duerksen, J. A., 1947: Manual of Geodetic Astronomy, Special Publication No. 237, U. S. Coast and Geodetic Survey (now National Ocean Survey, NOAA), Dept. of Commerce, Washington, D. C., 206 pp (NTIS Accession No. PB 267465).

King, R. G., Counselman, C. C., Shapiro, I. I., 1978: "Universal Time: Results from Lunar Laser Ranging," Journal of Geophysical Research, Vol. 83, No. 37, pp. 3377-3381.

Kolenkiewicz, R., Smith, D. E. Rubincam, D. P., Dunn, P. J., Torrence, M. H., 1977: "Polar Motion and Earth Tides from Laser Tracking," Philosophical Transactions Royal Society of London 284, pp. 485-494.

Newcomb, S., 1891: "On the Periodic Variation of Latitude, and the Observations with the Washington Prime-Vertical Transit," The Astronomical Journal No. 251, pp. 81-83.

Oesterwinter, C., 1978: "Polar Motion through 1977 from Doppler Satellite Observations," Proceedings of International Astronomical Union Symposium No. 82, "Time and the Earth's Rotation," Cadiz, Spain, May 1978.

Robertson, D. S., Carter, W. E., Corey, B. E., Cotton, W. D., Counselman, C. C., Shapiro, I. I., Wittels, J. J., Hinteregger, H. F., Knight, C. A., Rogers, A. E. E., Whitney, A. R., Ryan, J. W., Clark, T. A., Coates, R. J., Ma, C., Moran, J. M., 1978: "Recent Results of Radio Interferometric Determinations of a Transcontinental Baseline, Polar Motion and Earth Rotation," Proceedings of International Astronomical Union Symposium No. 82, Time and the Earth's Rotation," Cadiz, Spain, May 1978.

Silverberg, E., 1978: "On the Effective Usage of Lunar Ranging for the Determination of the Earth's Orientation," Proceedings IAU Symposium No. 82, Time and the Earth's Rotation, Cadiz, Spain, May 1978.

Smith, D. E., 1978: "Polar Motion from Laser Tracking Data," Goddard Space Flight Center.

Smith, D. E., Kolenkiewicz, R., Dunn, P. J., Torrence, M., 1978: "Determination of Polar Motion and Earth Rotation from Laser Tracking of Satellites," Proceedings IAU Symposium No. 82, Time and the Earth's Rotation, Cadiz, Spain, May 1978.

Smylie, D. E. and Mansinha, L., 1968: "Earthquakes and the Observed Motion of the Rotation Pole," Journal Geophysical Research 73, pp. 7661-7673.

Stolz, A., Bender, P. L., Faller, J. E., Silverberg, E. C., Mulholland, J. D., Shelus, P. J., Williams, J. G., Carter, W. E., Currie, D. G., Kaula, W. M., 1976: "Earth Rotation Measured by Lunar Laser Ranging," Science Vol. 193, pp. 997-999.

Yumi, S., 1964: Annual Report of the International Polar Motion Service for the Year 1962.

Questions and Answers

DR. TOM CLARK, NASA Goddard Space Flight Center:

Do we have any questions or comments? I might make one comment regarding Bill's last picture. In addition to the three stations he showed there, the same group is also conducting even earlier prototype observations using a Mark III system personally funded by DMA that will be going into Sweden, and some hardware personally funded by NASA which will be going into a station in California, so there will actually be some FY 79 observations going on at those stations.

Page intentionally left blank

TWO-WAY TIME TRANSFERS BETWEEN NRC/NBS AND NRC/USNO
VIA THE HERMES (CTS) SATELLITE

C.C. Costain, H. Daams, and J.S. Boulanger
National Research Council, Time and Frequency Section
Ottawa, Ontario, Canada

D. W. Hanson
National Bureau of Standards, Time and Frequency Division
Boulder, Colorado

W. J. Klepczynski
U. S. Naval Observatory, Time Service Division
Washington, D. C.

ABSTRACT

Two-way time transfers via the Hermes (CTS) satellite between NRC, Ottawa and NBS, Boulder and between NRC and USNO, Washington, D.C., began once a week in July 1978. At each station the differences were measured between the local UTC seconds pulse and the remote UTC pulse received by satellite. The difference between the readings, if station delays are assumed to be symmetrical, is two times the difference between the clocks at the two ground station sites. Over a 20-minute period, the precision over the satellite is better than 1 ns. The time transfer from NRC to the CRC satellite terminal near Ottawa and from NBS to the Denver HEW terminal are still subject to larger uncertainties which are being examined. The absolute measure of UTC differences also depends on the measurement of station delays, which in present circumstances will be difficult to carry out.

Two years ago, discussions were started on the possibility of two-way time transfer between the NRC and LPTF in France using the Symphonie satellite. A year ago, the possibility of participating in the third year of experiments on the Canadian-USA Hermes satellite arose, and the NRC applied to the Canadian Committee for an allocation of 1/2 hour a week. This experiment received final approval by the Minister of the Canadian Department of Communications in April 1978. The approved schedule, with NBS Boulder as partner, was for 16:00-16:30 ET on Fridays to September 31, and on Thursdays until December 31. The usage of the satellite alternates daily between Canada and the USA. Early in the experiment, the USNO was included, and in October an additional 1/2 hour in the Canadian schedule was allocated, 12:00-12:30 ET on Tuesdays, for time transfer between NRC/USNO.

It is perhaps necessary to state that the very short lead time, the shortage of manpower and equipment, and the operation through host terminals, has for all three laboratories placed constraints on the program.

Figure 1 is a schematic of the mini network. It shows a) the NRC/NBS link, and b) the NRC/USNO link. It was intended at one time to operate a USNO/NBS link, but this project was abandoned for technical reasons which will be evident later.

The Hermes satellite operates in the 12-14 GHz bands, and has two transmitters, one 200 W and one 20 W. Each channel has a steerable 2° beam which can be boresighted on a particular station as desired. The beams cannot normally be changed during an experiment, particularly since the failing telemetry on Hermes requires a large NASA antenna for acquisition of telemetry, and telemetry status of the satellite is rarely available these days.

All satellite terminals use a standard TV channel which requires a 1 volt peak-to-peak video input. The 1 pps and the TV horizontal sync pulses simulate the TV video and maintain proper video levels. It is realized that using a 1 pps is not the most efficient way to use the 6 MHz bandwidth, but NRC was already committed to using this format with France. It is also an easy format to put into operation quickly since all have 1 pps clock pulses available.

Because the NRC uses the master terminal at the Communications Research Center (CRC) with a 30-foot antenna, the 20 W beam is directed to Ottawa, and the 200 W to Denver or to Washington.

In Denver, NBS has the use of a terminal owned and operated by the Department of Health, Education and Welfare (HEW). It has an 8-foot antenna and 400 W transmitter power. This gives, at CRC in Ottawa, a received carrier of 40 db S/N with a 300 kHz bandwidth in the television channel.

In Washington the story is more complicated. In late July, experiments started with USNO using a NASA terminal at Goddard, but unfortunately we had no success. Initially the boresight was directed not to Ottawa, but halfway between Ottawa and Goddard. The signal to noise, down about 3 db, seemed adequate, but a later slide will show some of the difficulties. Trials began again some seven weeks later, after NASA had been off the air to make changes at the site, but the experiments were not successful. The additional time allocation was used, but NRC and USNO computer failure and interface problems compounded to frustrate the experiment.

At the Hermes Users Meeting in Wingspread, Wisconsin, September 19-21, 1978, Mr. K. Kaiser suggested that it might be possible to put a COMSAT Hermes terminal at the USNO. Since this had the enormous

advantage of avoiding the requirement for a time transfer from the laboratory to the terminal, and saves so much time in travel and equipment displacement, this was followed up with some enthusiasm by the USNO, and the COMSAT terminal was placed at the USNO. Successful transfers for the past three weeks have been achieved.

The time transfer from NBS to the HEW Denver terminal has been done by carrying two Cs standards to the terminal (Fig. 2), using one as the station clock, and obtaining closure after the experiment at NBS (Fig. 3). This has given 1-2 ns accuracy, and has shown that there are large and variable errors in the new TV Line-10 receivers.

In Ottawa, the NRC has been relying on the TV Line-10 receivers. CRC is about 32 kilometers (20 miles) from NRC, and both sites have a clear line of sight to the TV transmitter. The path difference has been indexed with portable clocks. Often, 1 ns standard deviation for 300 readings is obtained, but variations over 200 ns have been seen. It is usually possible to tell when the readings are reliable, but there is no doubt at the moment that the NRC/CRC link gives the largest source of error in the transfer, and NRC will convert to the NBS type of operation as soon as possible.

The principle of the measurements is shown in Fig. 4. At each station, the counter is started with the local UTC pulse and stopped with the pulse received from the satellite. The transit times are about 0.255 s for CRC/Denver, and 0.258 s for CRC/USNO. To obtain the difference between the two station clocks, the readings T_1 and T_2 for each second must be subtracted and divided by two. In fact, it is preferable to fit a cubic equation to the 1000-odd measurements and then exchange and subtract the equations.

There are the two other terms in the equation, hopefully small and constant. The terms t_1 and r_1 are the delays in the transmitter and receiver at station 1, and t_2 and r_2 the delays at station 2. It is fairly easy to measure the sum of t_1 and r_1, and this is done daily at some stations as a check on the constancy. To measure one or both individually is difficult, and this will be discussed later.

Figure 5 is a typical result from an early run in Ottawa. Part of the plot of the data-equation is given and also a histogram of the difference between the data and the equation. It is not a gaussian, but by selecting the gate for the data, it is possible to center the apparent gaussian at zero offset.

However, this same day NBS reported a standard deviation of 1.4 ns on all points, and NRC should have an advantage in signal to noise, with a factor of 25 in antenna gain to offset the factor of 0.10 in power. At this point, suspicions arose of the off-boresight operation (with the boresight between Ottawa and Goddard) even though the signal

was less than 3 db down. The boresight was returned to Ottawa. The result is shown in Fig. 6 with a standard deviation of 1.4 ns for 1091 points. Needless to say, the experiment has been run since that time only with the antennas boresighted on the stations.

The degradation of the pulse off the boresight was puzzling, and satellite engineers found the effect difficult to believe. However, in Fig. 7, the worst example shows the reception in Ottawa with the antenna boresight on Washington, where most of the pulses appear to arrive 150 ns late.

Consequently, a three-hour experiment was set up in Ottawa for a loop test through the satellite, with various off-boresight settings of both satellite and ground antenna. No effects were seen except for an increased standard deviation as the signal to noise decreased, and no delayed pulses distorted the gaussian.

There was only one change made in the CRC terminal facilities between the two experiments, and it was significant. The NRC equipment is housed in the Symphonie terminal at CRC, and for the Hermes experiment, the pulses are sent and received over 2.4 km (1.5 mi) of triax cables to the Hermes terminal. Before the experiment started, the cable terminations or equalizers were very carefully adjusted to reproduce or reconstruct the 1 pps signal. Unknown to NRC, a severe lightning storm burned out those equalizers which were, of course, replaced. The new equalizers overemphasize the high frequencies; in fact, there is an overshoot and ringing on both the rise and fall of the pulse. Because the results are better, we have not requested adjustment of the equalizers.

It still appears that there was a serious degradation of the pulse, a loss of high frequencies when off-axis which has now been over-compensated. But the question must be settled, and it is intended to carry out an experiment with the full bandwidth of 70 MHz of Hermes, with 20 ns pulses and a high repetition rate, to see if diffraction and refraction do cause pulse degradation off-axis of the beam. If this does occur, and some say it must, it will place a serious limitation on the bit rate that can be transmitted off the boresight of any beam.

The results of the time transfer between NRC and NBS are given in Fig. 8. It must be emphasized that no measurements and no corrections have been made for terminal delays, and therefore all quoted values for the differences in the UTC(i) obtained by satellite must be corrected by the portable clock results. The standard deviation of the points from the two straight lines is 9 ns, a figure that is inclined to inspire confidence. It was immediately assumed at NRC that the change of 1.5×10^{-13} was the result of evaluation of CsV on September 15-16, even though the change in frequency was three times the expected error.

However, there is now evidence to show that CsV did not change in frequency by this amount. There is no evidence of a change in the rate in the direct UTC(NRC) - UTC(USNO) transfer via Loran-C. Further, a change of 1.5×10^{-13} would have resulted in a 0.5 μs error in the USNO portable clock value on October 31, 1978.

NBS evaluated their NBS-6 primary cesium standard around the end of October 1978 and obtained a normalized frequency difference in UTC(NBS) - NBS-6 (sea level) of 0.6×10^{-13}. Over this period, from Fig. 8, the normalized frequency difference of the two time scales UTC(NRC) - UTC(NBS) was 1×10^{-13}. The frequency of UTC(NRC) is the frequency of CsV corrected to sea level, and therefore the normalized frequency difference for the primary standards CsV (sea level) and NBS-6 (sea level) was 1.6×10^{-13}. This is the same order as the difference observed over the past few years, which indicates that there was not a large change in the CsV frequency on September 15, 1978.

Some uncertainty was introduced by the changes at CRC that were made necessary by the lightning damage during the week of August 25, 1978. However, the internal consistency of the satellite time transfer is such that it is concluded that changes in one or probably both time scales produced the observed 1.5×10^{-13} change in the normalized frequency difference.

The NRC/USNO preliminary results for November 14, 21, and 28, are respectively 4938, 5033, and 5053 ns. These give a frequency difference similar to the NRC/NBS results for the same period.

The experiment has not yet quite matched the precision achieved by Chi (1975) and Saburi (1976) in experiments by NASA and the Radio Research Laboratory of Japan, but the main virtue of this experiment is the transfer over a long period between very stable time scales. It has been exciting in the potential that it shows, and instructive in the weaknesses that have become evident in the present setup. In an operational mode, the following three steps should yield 1 ns precision and an accuracy of a few nanoseconds:

1. The acquisition of matched PRN code generators by all participants. The 1 pps system used in this experiment is too vulnerable to changes in transmission, trigger levels, etc, and the advantage of averaging over many pulses is lost.

2. The acquisition by all participants of "private" terminals to operate at the laboratories. It is essential to eliminate the errors in the transfer of time from the laboratory to the terminal. It is also essential to measure and monitor the terminal delays, which is very difficult to do when the "host" terminal is being used for many experiments.

3. A fully automatic computerized operation will be necessary to operate in short burst during the night to take advantage of unused satellite time.

An experiment has been approved for the Anik B satellite which is expected to be launched in December 1978. This is primarily a VLBI experiment between the NRC radio astronomy stations in Penticton, British Columbia, Algonquin Park, Ontario, and the Naval Research Laboratory station in Maryland. It is intended to maintain coherence among the local oscillators at these stations by exchanging side tones through the Anik B satellite. Professor Yen of the University of Toronto is predicting 10 ps precision for this experiment, and the NRC Time Laboratory has been invited to join with a fourth station in Ottawa as a time station only. It is hoped that this will be possible and that a 12 - 14 GHz terminal can be acquired for this purpose as the first stage of the three requirements given earlier.

There are not many results to report as yet from the NRC/LPTF time transfer via the Symphonie satellite. This is to run from July 1, 1978 to July 1, 1979 on a daily basis, except for outages of about 6 weeks during eclipse periods. The data are being processed in France, except for some 30 readings exchanged daily by the voice channel. They expect to improve their data handling capacity in the near future. The terminal used in Ottawa, operated by NRC, has a 30-foot antenna. The France terminal is the main control center at Pleumeur Bodou in Brittany. There is an uncertainty of about 30 ns in the time transfer via the TV network to Paris. The preliminary figures give a normalized frequency difference between UTC(NRC) and UTC(OP) of $1 \times 10^{-13} \pm 0.2 \times 10^{-13}$ for July and August 1978.

In support of the time transfer, several portable clock trips have been made. In particular, in the last week of October, there were successful clock trips to NRC from Observatoire de Paris, USNO and NBS.

We must acknowledge a great deal of assistance in carrying out this experiment. The NRC has had the use of the CRC terminal and the fullest support from the Hermes group at CRC and others of the Department of Communication, from the experiment controller John Brookfield, the engineers, programmers and scientists. The NBS is pleased to acknowledge the weekly loan of the Denver terminal of HEW, the assistance of Mr. Earl Henderson, National Library of Medicine in Bethesda, Maryland, and of Mr. Bernie Lackey, who operates the Denver terminal so efficiently. The USNO has had the cooperation of NASA Goddard, and now of Kim Kaiser and Wes Venstra of COMSAT with their terminal at the USNO.

REFERENCES

1. Chi, A. R., and Byron, E., "A Two-Way Time Transfer Experiment Using a Synchronous Satellite," Proc. of the 7th Precise Time and Time Interval (PTTI) Applications and Planning Meeting, pp. 357-376 (December 1975).

2. Saburi, Yoshikazu, Yamamoto, Minoru, and Harada, Kikuo, "High-Precision Time Comparison via Satellite and Observed Discrepancy of Synchronization," IEEE Trans. Instr. & Meas., IM-25, No. 4, pp. 473-477 (December 1976).

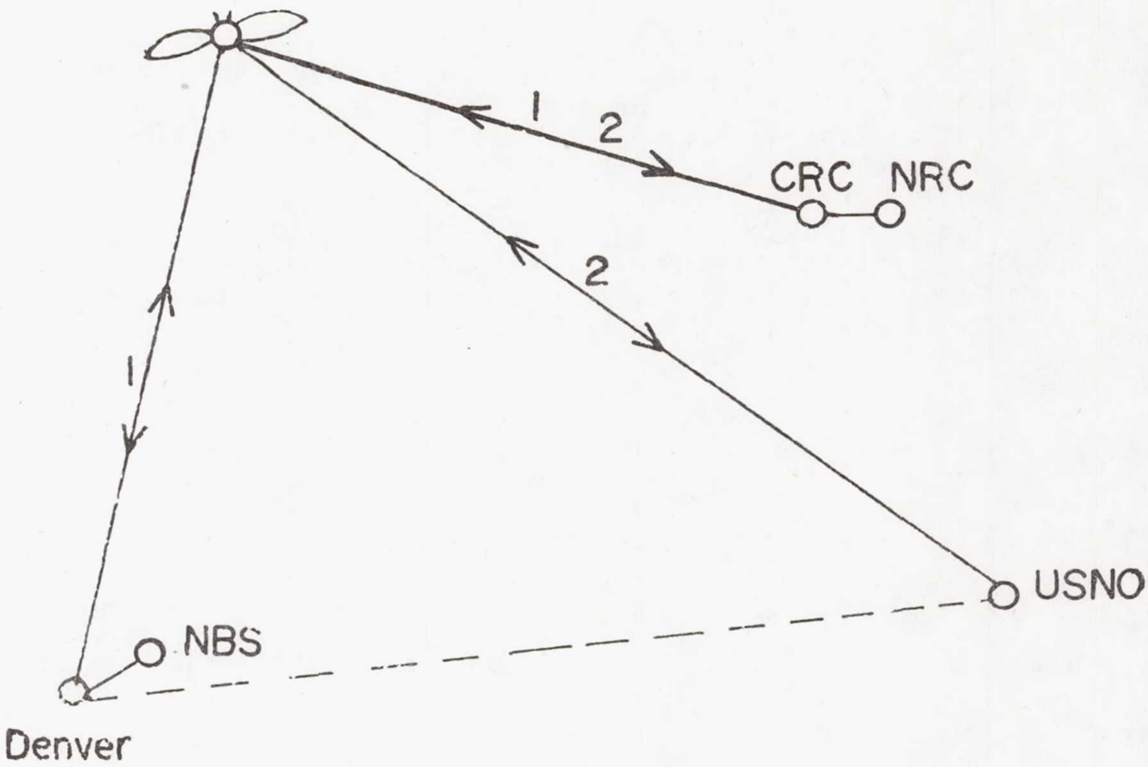


FIG. 1a. TWO-WAY TIME TRANSFER PATHS FOR NRC/NBS AND NRC/USNO USING THE HERMES (CTS) GEOSTATIONARY SATELLITE.

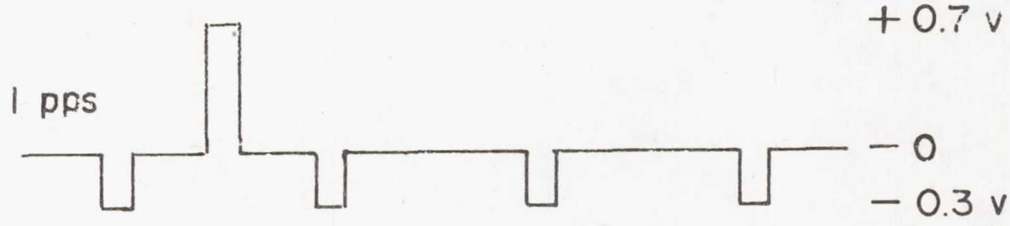


FIG. 1b. VIDEO FORMAT WITH 1 pps TIME PULSES AND THE 15,625 Hz TV HORIZONTAL SYNC PULSES.

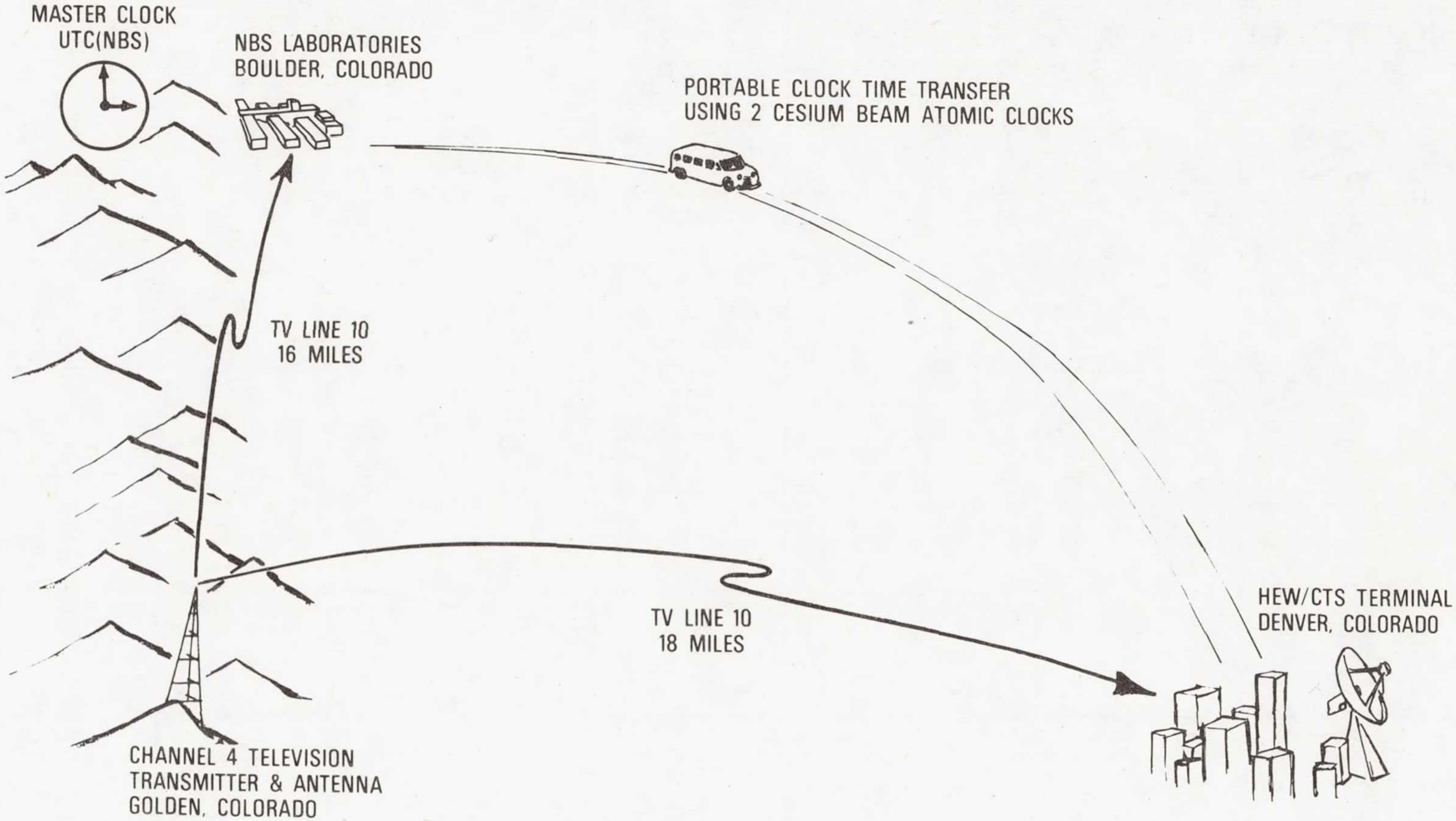


FIG. 2. TIME TRANSFER FROM NBS BOULDER TO THE HEW TERMINAL, DENVER, VIA PORTABLE Cs CLOCKS AND TV LINE 10.

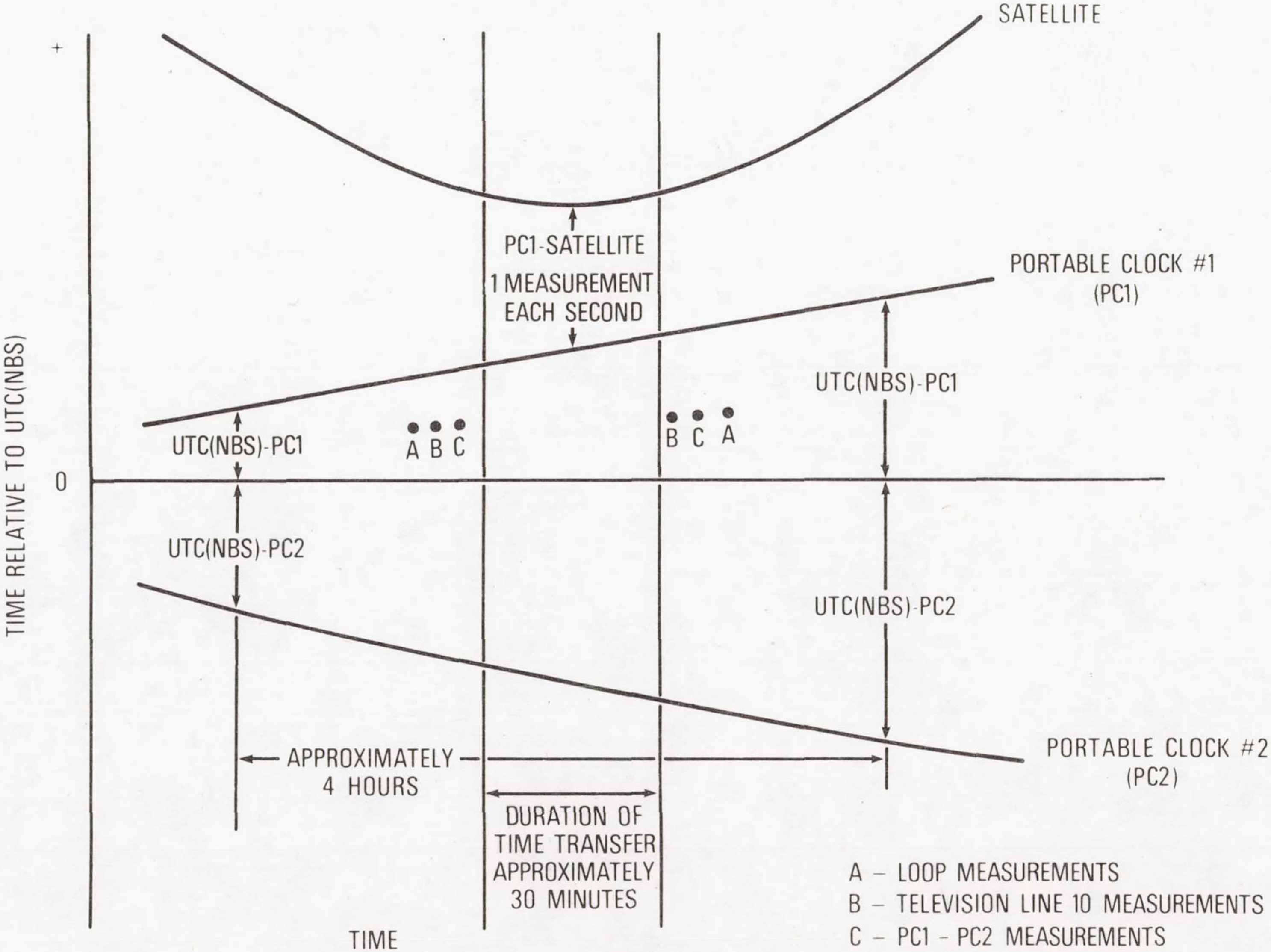


FIG. 3. DETAILS OF THE NBS/HEW TIME TRANSFER.

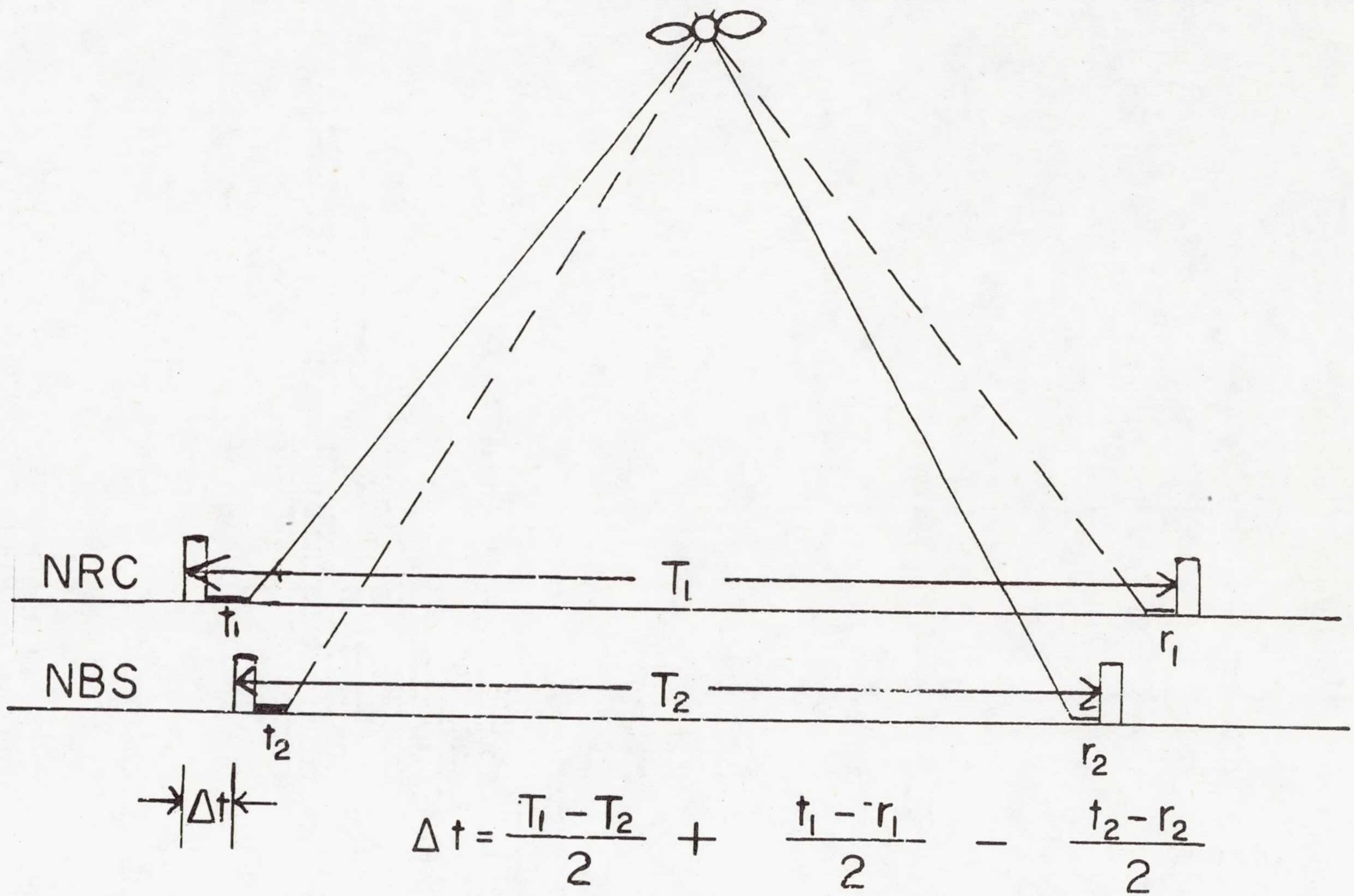


FIG. 4. TWO-WAY SATELLITE TIME TRANSFER MEASUREMENTS.


```
Histogram : scale =  1:1,computed from 195550 to 201321
 AT >24 :    40 |****************************************
 AT  24 :     2 |**
 AT  23 :     3 |***
 AT  22 :     2 |**
 AT  21 :     4 |****
 AT  20 :     3 |***
 AT  19 :     9 |*********
 AT  18 :     7 |*******
 AT  17 :     9 |*********
 AT  16 :     4 |****
 AT  15 :    16 |****************
 AT  14 :    17 |*****************
 AT  13 :    22 |**********************
 AT  12 :    26 |**************************
 AT  11 :    21 |*********************
 AT  10 :    21 |*********************
 AT   9 :    42 |******************************************
 AT   8 :    44 |********************************************
 AT   7 :    57 |*********************************************************
 AT   6 :    55 |*******************************************************
 AT   5 :    56 |********************************************************
 AT   4 :    68 |********************************************************************
 AT   3 :    62 |**************************************************************
 AT   2 :    66 |******************************************************************
 AT   1 :    63 |***************************************************************
 AT   0 :    87 |***************************************************************************************
 AT  -1 :    79 |*******************************************************************************
 AT  -2 :    70 |**********************************************************************
 AT  -3 :    43 |*******************************************
 AT  -4 :    35 |***********************************
 AT  -5 :    10 |**********
 AT  -6 :     7 |*******
 AT  -7 :     1 |*
 AT<-24 :     1 |*

The file is sp230195548
Time index is 200000
 Y(t)(ns) =  255594062.396 -2.1353568194363E+01*t +1.8252224804267E-03*t*t +1.6942143317262E-08*t*t*t
Computed from 195550 to 201321
 Standard deviation =        1.401 in a window of         5 ns.
   687 points discarded on a total of  1052.
```

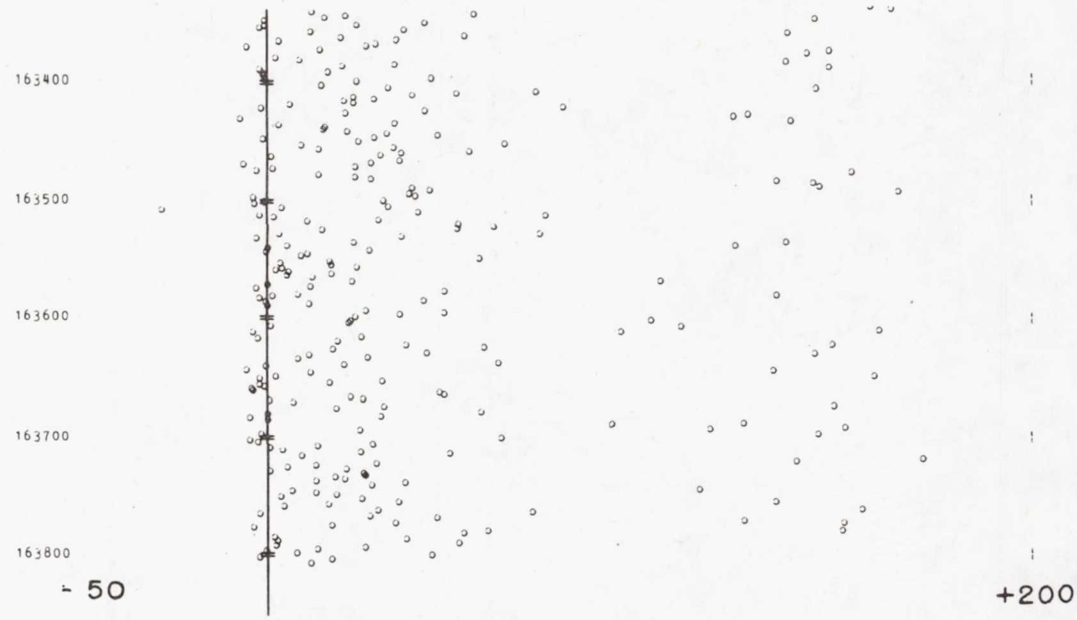
FIG. 5. NRC/NBS TRANSFER. AN NRC PLOT AND HISTOGRAM OF DATA-EQUATION, IN ns, WITH ANTENNA BORESIGHT BETWEEN OTTAWA AND WASHINGTON.

Histogram : scale = 3:1,computed from 201149 to 203001

AT 16 :	1
AT 15 :	0
AT 14 :	0
AT 13 :	0
AT 12 :	0
AT 11 :	0
AT 10 :	0
AT 9 :	0
AT 8 :	0
AT 7 :	0
AT 6 :	0
AT 5 :	0
AT 4 :	10
AT 3 :	28
AT 2 :	125
AT 1 :	232
AT 0 :	292
AT -1 :	243
AT -2 :	116
AT -3 :	39
AT -4 :	6
AT -5 :	0
AT -6 :	0
AT -7 :	0
AT -8 :	0
AT -9 :	0
AT -10 :	0
AT -11 :	0
AT -12 :	0
AT -13 :	0
AT -14 :	0
AT -15 :	0
AT -16 :	0
AT -17 :	0
AT -18 :	0
AT -19 :	0
AT -20 :	0
AT -21 :	0
AT -22 :	1

```
The file is sp244201146
Time index is 200000
 Y(t)(ns) =  255532665.594 -1.3587399250907E+01*t +1.9640465177748E-03*t*t +4.8304808280294E-09*t*t*t
Computed from 201149 to 203001
 Standard deviation =        1.418 in a window of        20 ns.
     2 points discarded on a total of  1093.
```

FIG. 6. NRC/NBS TRANSFER. AN NRC PLOT AND HISTOGRAM OF DATA-EQUATION, IN ns, WITH ANTENNA BORESIGHT ON OTTAWA.




```
Histogram : scale =  1:1,computed from 162802 to 163827
 AT >24 :   254 |***********************************************************************************************************
 AT  24 :     9 |*********
 AT  23 :    12 |************
 AT  22 :     8 |********
 AT  21 :     8 |********
 AT  20 :     9 |*********
 AT  19 :     6 |******
 AT  18 :     6 |******
 AT  17 :    11 |***********
 AT  16 :     8 |********
 AT  15 :     6 |******
 AT  14 :     8 |********
 AT  13 :    12 |************
 AT  12 :    13 |*************
 AT  11 :     6 |******
 AT  10 :     5 |*****
 AT   9 :    11 |***********
 AT   8 :     9 |*********
 AT   7 :     6 |******
 AT   6 :    11 |***********
 AT   5 :    10 |**********
 AT   4 :     8 |********
 AT   3 :    11 |***********
 AT   2 :    12 |************
 AT   1 :    14 |**************
 AT   0 :    19 |*******************
 AT  -1 :    24 |************************
 AT  -2 :    19 |*******************
 AT  -3 :    10 |**********
 AT  -4 :    11 |***********
 AT  -5 :     4 |****
 AT  -6 :     5 |*****
 AT  -7 :     3 |***
 AT  -8 :     0 |
 AT  -9 :     0 |
 AT -10 :     0 |
 AT -11 :     0 |
 AT -12 :     0 |
 AT -13 :     0 |
 AT -14 :     0 |
 AT -15 :     0 |
 AT -16 :     0 |
 AT -17 :     0 |
 AT -18 :     0 |
 AT -19 :     1 |*

The file is sp216162800
Time index is 162800
 Y(t)(ns) =  260832707.579 -5.0377853871081E+01*t -5.8989750437129E-05*t*t +8.8965515574784E-08*t*t*t
Computed from 162802 to 163827
 Standard deviation =        2.591 in a window of        10 ns.
   489 points discarded on a total of   626.
```

FIG. 7. NRC/NRC LOOP TEST. AN NRC PLOT AND HISTOGRAM OF DATA-EQUATION, IN ns, WITH ANTENNA BORESIGHT ON WASHINGTON.

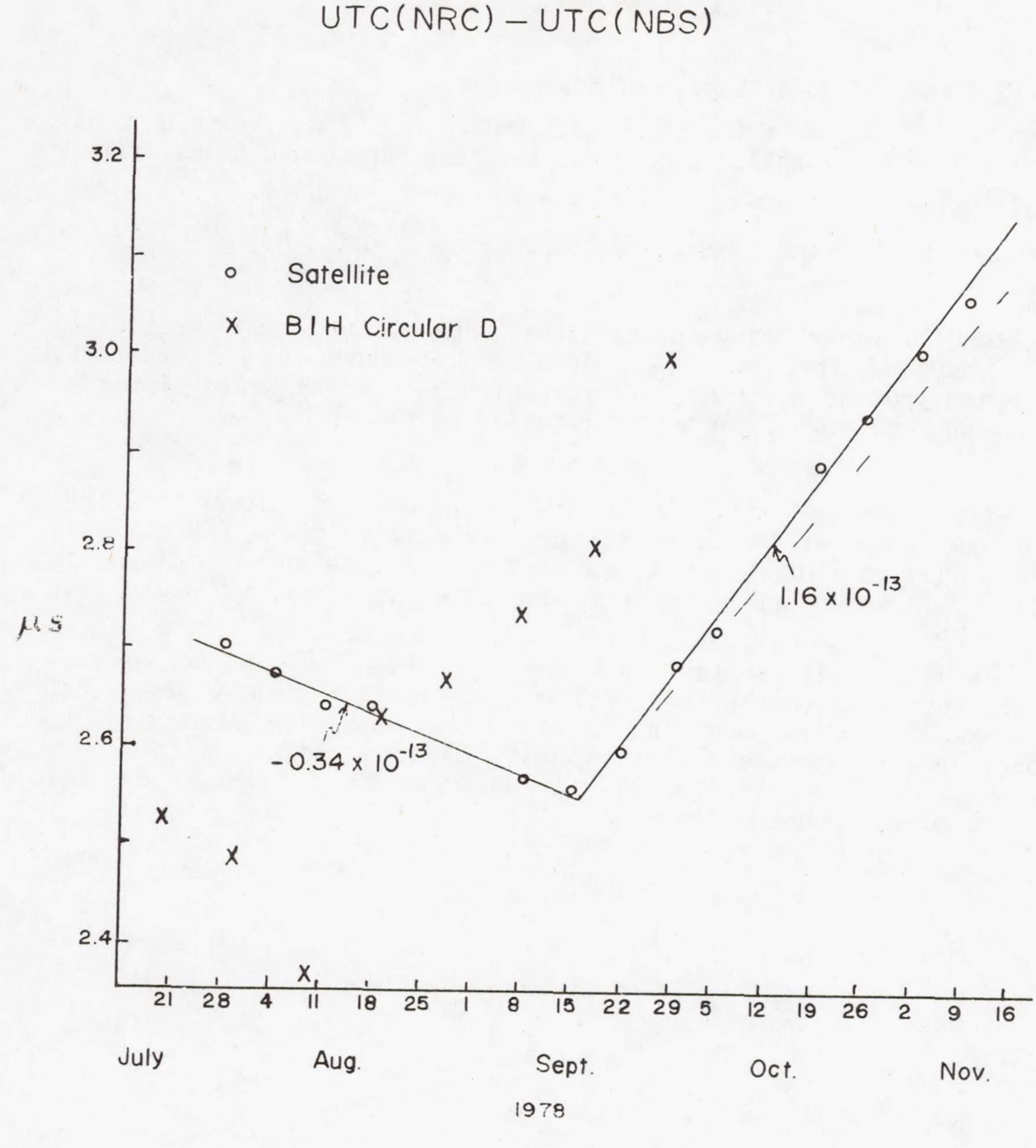


FIG. 8. UTC(NRC) - UTC(NBS) IN MICROSECONDS. THE CIRCLES WERE OBTAINED FROM THE SATELLITE TIME TRANSFER, THE CROSSES FROM BIH CIRCULAR D. THE DIFFERENCE IN SLOPE BETWEEN THE SOLID LINE AND DASHED LINE IS 1×10^{-14}.

Questions and Answers

MR. DAVID ALLAN, National Bureau of Standards:

I would like to make two comments: Number one, I'd like to give Dr. Barnes credit for having picked out that one nanosecond granularity.

DR. COSTAIN:

Belanger said it was. He couldn't convince me until May.

MR. ALLAN:

Right. The other thing is that it is interesting to note that using the equations that you wrote down, if you compare your numbers, assuming you had accuracy, via portable clock, they would differ by about 60 nanoseconds due to the rotation of the earth.

DR. COSTAIN:

I did not mention that. Certainly, if one moves to higher accuracy, that goes with the measuring of the terminal. I think you all said the figure of about eight nanoseconds to Washington and 160 to France, or something. And if we are after accuracy, we would have to do that.

I hope it is constant, although I remember in Saburi's experiment, the variation of the satellite, in fact, with his precision, was enough to see that in the variation in path. The satellite moves in our experiment. It is just going to drift a little more eliptical all the time. We get about up to 50 nanoseconds per second movement now, I think.

SUBMICROSECOND COMPARISONS OF TIME STANDARDS VIA THE NAVIGATION TECHNOLOGY SATELLITES (NTS)

J. Buisson, T. McCaskill, J. Oaks, and D. Lynch
U. S. Naval Research Laboratory, Washington, D. C.
C. Wardrip, NASA Goddard Space Flight Center
G. Whitworth, Applied Physics Laboratory

ABSTRACT

During May through September 1978 a six nation cooperative experiment was performed to intercompare time standards of major laboratories at the submicrosecond level using NTS satellites.

NTS time transfer receivers, which were developed for use with the NTS series of satellites were installed at the Division of National Mapping (DNM), Australia; National Research Council (NRC), Canada; Royal Greenwich Observatory (RGO), England; Bureau International de l'Heure, France (BIH); Institute for Applied Geodesy (IFAG), West Germany; and in the U.S. at the Goddard Space Flight Center (GSFC), National Bureau of Standards (NBS), Naval Research Laboratory (NRL) and the Naval Observatory (USNO).

The results of the clock intercomparisons will be presented.

INTRODUCTION

The major objective was to perform an interim demonstration of the time transfer capability of the NAVSTAR GPS system using a single NTS satellite. Measurements of time difference (pseudo-range) are made from the NTS tracking network and at the participating observatories. The NTS network measurements are used to compute the NTS orbit trajectory. The central NTS tracking station has a time link to the Naval Observatory UTC(USNO,MC1) master clock. Using measurements taken with the NTS receiver at the remote observatory, the time transfer value UTC(USNO,MC1)-UTC(REMOTE, VIA NTS) is calculated. For GPS, a similar procedure could be followed using simultaneous measurements taken between the user and four GPS satellites. With the four GPS pseudo-range (time difference)

measurements taken at an (unknown) location the user may solve for three position coordinates in addition to time offset with respect to GPS time. The goal for the NTS effort was to achieve worldwide time transfer of less than one microsecond accuracy.

A second objective was to compute weekly worldwide intercomparisons of the observatory clock offsets using predicted values of satellite clock offset and ephemeris. Each participant enters appropriate measurements into computer files which are later processed. Other objectives include co-location at laser sites and the use of the observatory time scales in evaluating the spacecraft clock performance.

Time Difference Measurements

Time difference (pseudo-range) measurements are made between the spacecraft and the user by side tone ranging (1). The NTS-2 spacecraft also has a GPS pseudo-random sequence transmitter. All measurements presented in this paper were made using the side tone ranging system, which has a resolution of 1.56 nsec (48 cm.). Measurements of time difference may be converted to pseudo-range by multiplying by the speed of light in a vacuum. Units of time are used in this paper to facilitate comparisons with the PTTI community.

The time difference measurement is composed of the difference between the satellite clock and the user clock, plus satellite transmitter delays, propagation path delay, ionospheric delay, tropospheric delay, user antenna delay, cable and receiver delay. All of these factors must be measured or estimated. In addition to the above factors, the spacecraft clock is influenced by the relativistic frequency shift, magnetic fields, energetic particles, and small variations in temperature and drive level.

Receivers of two designs were employed in making the measurements. One receiver (2) made measurements at a nominal UHF frequency of 335 MHz. The second receiver used was capable of making measurements at the L band frequency of 1580 MHz in addition to the UHF frequency. The two channel receiver measurements were combined, by software, to correct for the first order ionospheric refraction.

Spacecraft Frequency Standards

Timing signals transmitted from NTS-2 are derived from a cesium frequency standard; NTS-1 employs both rubidium and quartz oscillators. Frequency stability results have been previously reported (3,4) for one of the NTS-2 cesium standards and for rubidium and quartz oscillators.

The NTS-2 cesium standard was used to measure the relativistic frequency shift (5) at the GPS constellation altitude. The NTS-2 cesium output frequency was adjusted so that the received frequency is near that of UTC(USNO,MC1). In contrast, the NTS-1 quartz oscillator is periodically adjusted in frequency and time. The maximum frequency excursions of the quartz varied from $+2 \times 10^{-9}$ to -4×10^{-9} with respect to UTC(USNO,MC1). Noteworthy is the fact that the ease of operation is superior with cesium, inasmuch as comparatively large periodic adjustments are required with the quartz frequency standard.

Time Transfer Technique

The time transfer to a remote location is obtained by four time links to UTC(USNO,MC1). The four links are (a) from the remote user clock to the spacecraft clock, (b) the spacecraft frequency time update for the time difference between observations obtained at the remote site and the central site, (c) from the central station clock to the spacecraft clock, and (d) from the central station clock to UTC(USNO,MC1). Figure (1) depicts the four links used in this procedure. This procedure incorporates the short to medium term stability of the spacecraft and control station clock with the long term stability of the U.S. Naval Observatory multi-clock time scale.

Measurements of [UTC(USNO,MC1)-UTC(REMOTE, VIA NTS)] may be taken with a variety of frequency sources of varying stability. The major observatories participating in this experiment possess frequency standards and time scales of proven accuracy, with sufficient difference in geographic location (figure 2), to check the time transfer at different positions of the spacecraft orbit.

Time Transfer Results

Figures (3)-(12) present time transfer results as determined from the NTS spacecraft. The figures are similar in format in as much as each remote observing station is referenced through the NTS central ground observing station located at Chesapeake Bay Division (CBD) of NRL. The CBD site is linked to the USNOMC by a series of portable clock closures to an accuracy of 10-20 nanoseconds.

Table 1 presents the phase offset and frequency difference of each remote station clock against the USNOMC for a given epoch time which is nominally placed in the middle of the observed data span. In addition, the RMS of a straight line least squares fit to all satellite passes observed by the remote station is presented as a measure of the noise in the time transfer values.

TABLE I

NTS
TIME TRANSFER RESULTS
UTC(USNO,MC)-UTC(REMOTE, NTS)

Remote Site	Epoch (day 1978)	Phase Offset (microsec)	Frequency (pp10^{11})	RMS (nanosec)
RGO (JP)	186	-160.734	-.245	369
BIH (OP)	156	1.574	-.006	318
CERGA	130	0.995	-.049	324
IFAG	186	- 10.277	-.017	377
DNM (590)	186	158.902	.329	458
RRL	303	- 18.050	-.010	862
NRLM	304	- 47.829	.003	998
NBS	151	- .474	-.014	398
NRC	186	- 3.716	-.005	152
USNO (MC1)	186	- .036	.000	171

From the table it can be seen that the two Japanese remote sites (RRL and NRLM) exhibit a higher noise level than the other observing stations. These higher noise level measurements were the result of using predicted satellite position ephemeris. Further analysis will be performed using observed orbital trajectory.

Also plotted in figures (3)-(12) are the results of portable clock closures performed by personnel from the USNO. These portable clock closures are used as "truth" or absolute accuracy tie-in for the NTS results.

Figures (13)-(15) present time transfer results from the NTS remote observing station located at the Panama Canal Zone (CZ) site. Results in this data span were obtained with both NTS2 and NTS1 spacecrafts. The NTS2 data included observations available at both 335 MHz and 1580 MHz, allowing for a first order ionospheric delay measurement.

The NTS1 measurements used single frequency measurement at 335 MHz. Table 2 summarizes the CZ results in a similar fashion to Table 1.

TABLE 2

NTS
TIME TRANSFER RESULTS
UTC(USNO,MC)-UTC(CZ)

Epoch (Day, 1978)	Phase (microsec)	Frequency ($pp10^{11}$)	RMS (nanosec)
119	-23.882	-.052	330
177	-26.293	-.093	63
170	-26.357	-.094	9

Figure (13) presents the entire data span consisting of both NTS2 and NTS1 measurements. Figure (14) presents only NTS2 data. The improvement in noise level was from 330 nanoseconds to 63 nanoseconds. This improvement was the results of two major advantages of the NTS2 spacecraft over the NTS1 spacecraft; firstly the use of a cesium oscillator in space (NTS2) as opposed to a quartz oscillator (NTS1) and, secondly, the ability to correct for the ionospheric delay by dual frequency measurement (NTS2).

The additional improvement in noise level between figures (14) and (15) (from 63 nanosec to 9 nanosec) is the result of a systematic effect in the orbit determination method which corresponds to the 2 rev/day orbit configuration. Figure (15) uses observations obtained from the same side of the orbit each day. This noise level of 9 nanoseconds is considered to be indicative of results which can be attained in the full operational GPS constellation.

System Closure

Figure (12) presents the time transfer results for a receiver located at the U. S. Naval Observatory with a direct input from UTC(USNO,MC1). It can be seen that the noise level is 171 nanosec with an offset of -36 nanosec at the epoch presented.

Time comparisons for five of the major observatories are presented in figure (16). The insert in figure (16) presents the offset of three of the observatories to permit relative frequency comparison.

Noteworthy is the line for UTC(USNO,MC1) via NTS; a small slope on the order of a few parts in 10(15) is present which is not statistically significant.

Table 3 presents the differences for the NTS1 time transfers with respect to the interpolated portable clock measurements. The average accuracy indicated by the portable clock is -0.06 usec. This table links the entire experiment to the absolute or "truth" values as determined by the DOD master clock.

TABLE 3

SUMMARY OF
PORTABLE CLOCK CLOSURES
VS
NTS TIME TRANSFER RESULTS

STATION	DAY (1978)	PORTABLE CLOCK-NTS TIME TRANSFER (US)
BIH	124	-.57
CERGA	117	.70
DNM	282	.09
IFAG	199	.03
NBS	221	.19
NRLM	299	-.53
RGO	115	.44
RRL	303	.13
USNO	186	.04

Conclusions

The following items are summarized as a conclusion for the six nation time transfer campaign:

o Time transfers via NTS satellites of better than 1 microsecond accuracy have been demonstrated.

- Simulated single satellite GPS operation has been demonstrated.

- A 9 nanosecond time transfer noise level over a 12 day span has been demonstrated as a possible best value of results.

Acknowledgements

Acknowledgement is given to the contributors of the NTS Data: Dr. G.M.R. Winkler, K. Putkovich and A. Johnson from the Naval Observatory (USNO), Washington, D. C., U.S.A.; D. W. Hanson, National Bureau of Standards (NBS), Boulder, Colorado, U.S.A.; Dr. C. C. Costain, National Research Council (NRC), Ottawa, Canada; Dr. B. Guinot, Bureau International de l'Heure (BIH), Paris, France; Dr. P. Morgan, Division of National Mapping (DNM), Queanbeyan, N.S.W. Australia; Dr. J. Pilkington, Royal Greenwich Observatory (RGO), East Sussex, England; Dr. J. Kovalevsky, Group de Recherches de Geodesie Spatiale (GRGS), Grasse, France; Dr. K. Nottarp, Institute fur Angewandte Geodasie (IFAG), Wettzell, Germany.

References

1. Landis, Paul G., Silverman, I., Weaver, Charles W., "A Navigation Technology Satellite Receiver", NRL Memorandum Report #3324, July 1976.

2. Raymond, L., Oaks, J., Osborne, J., Whitworth, G., Buisson, J., Landis, P., Wardrip, C., Perry, J., "Navigation Technology Satellite (NTS) Low Cost Timing Receiver Development", Proceedings of the Eighth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, November, 1976.

3. McCaskill, T. B., White, J. W., Stebbins, S., Buisson, J. A., "NTS-2 Cesium Frequency Stability Results", Proceedings of the 32nd Annual Symposium on Frequency Control, 1978.

4. McCaskill, T.B., and Buisson, J.A., "NTS-1 (TIMATION-III) Quartz and Rubidium Oscillator Frequency Stability Results", NRL Report 7932, December 12, 1975.

5. Buisson, J.A., Easton, R.L., McCaskill, T.B., "Initial Results of the NAVSTAR GPS NTS-2 Satellite", Proceedings of the Ninth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, March, 1978.

FIGURES

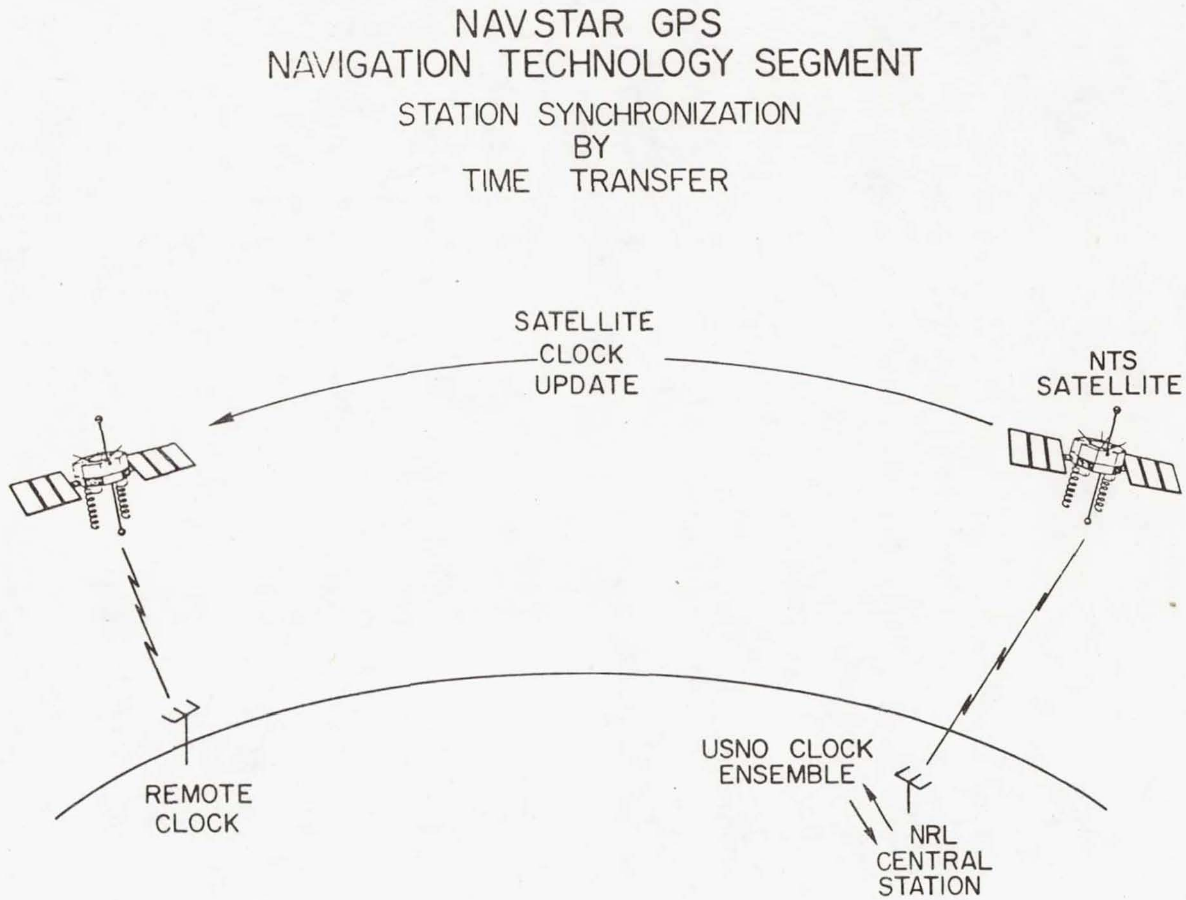


Figure 1 Time Transfer Configuration

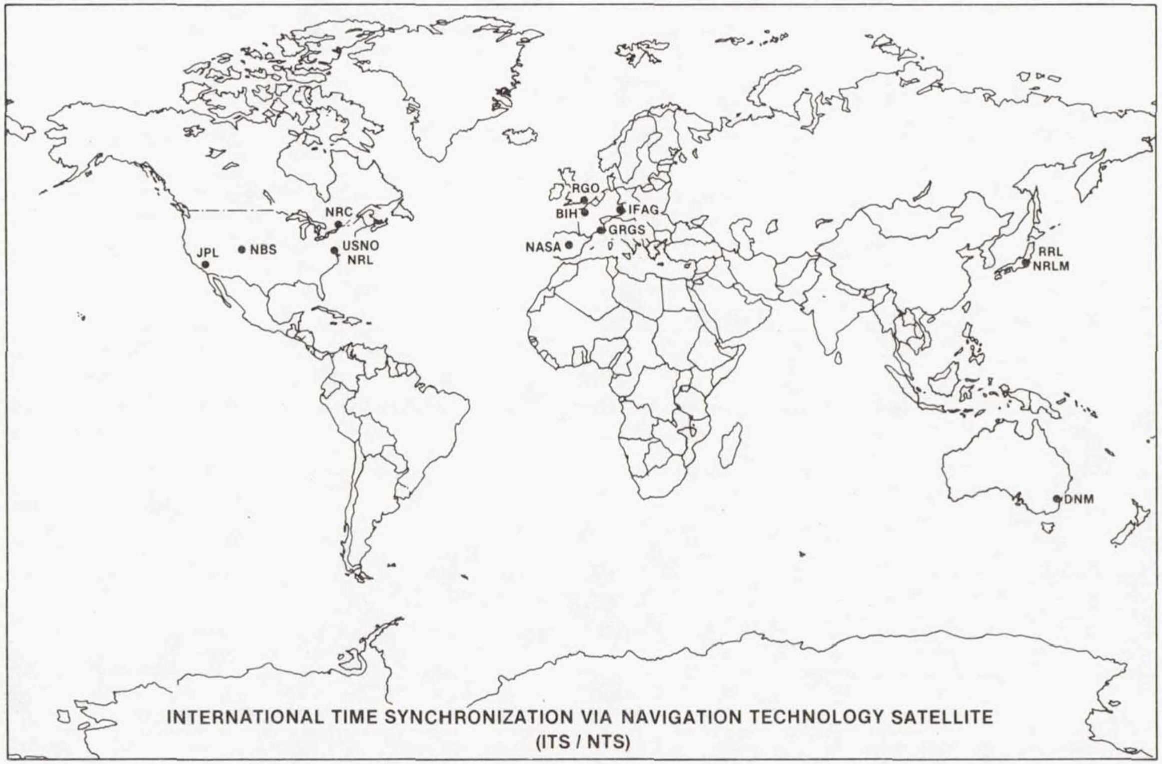


Figure 2 International Time Synchronization via Navigation Technology Satellite

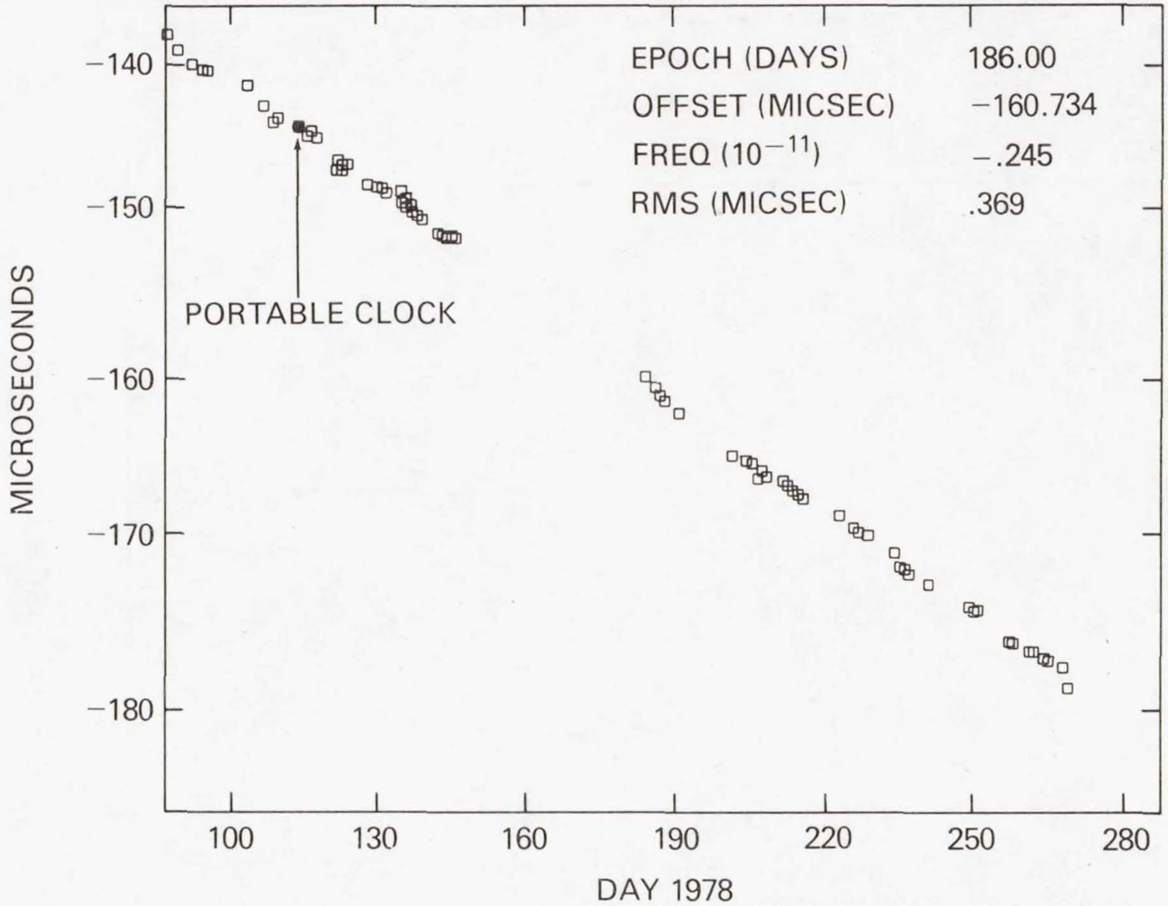


Figure 3 Time Transfer Results from Royal Greenwich Observatory (USNO,MC1)-(RGO,JP)

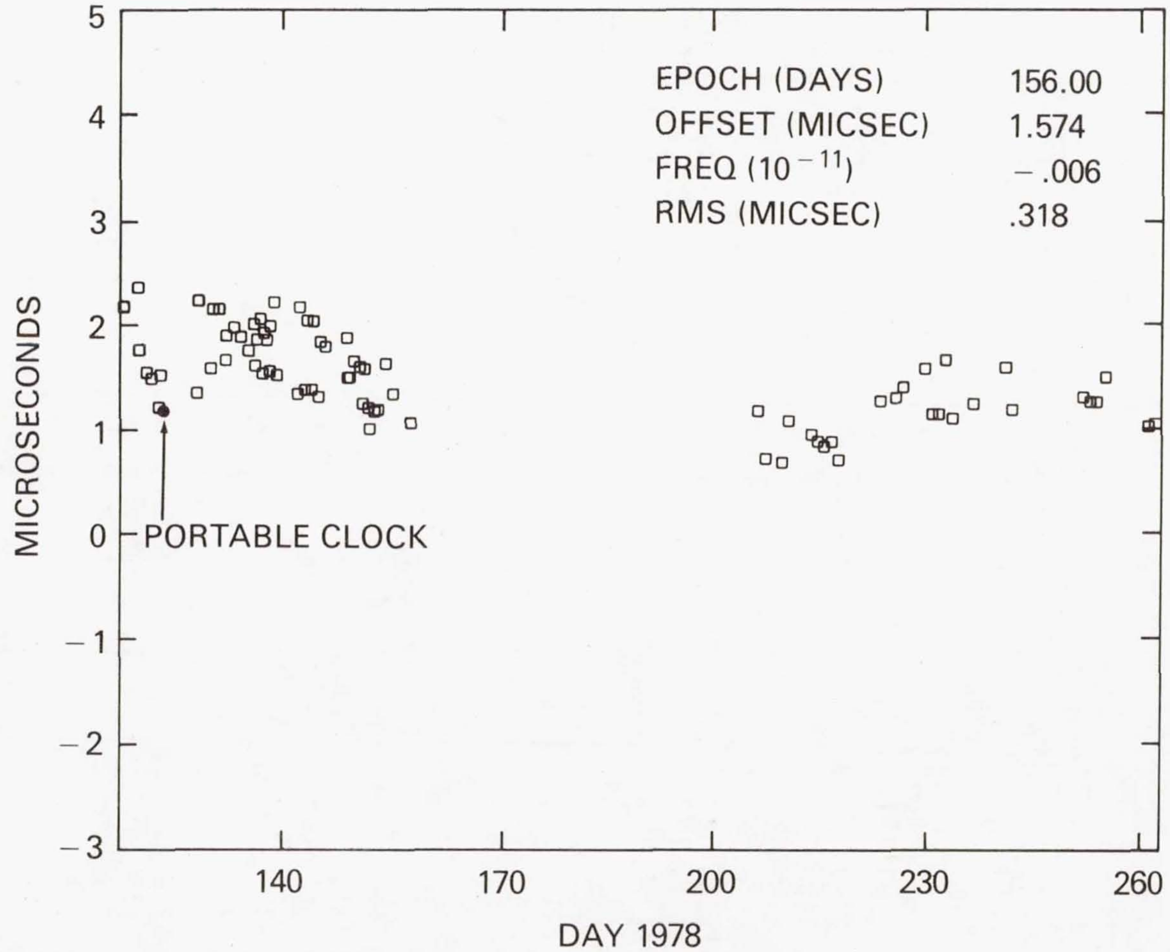


Figure 4 Time Transfer Results from Paris OP
(USNO,MC1)-(OP)

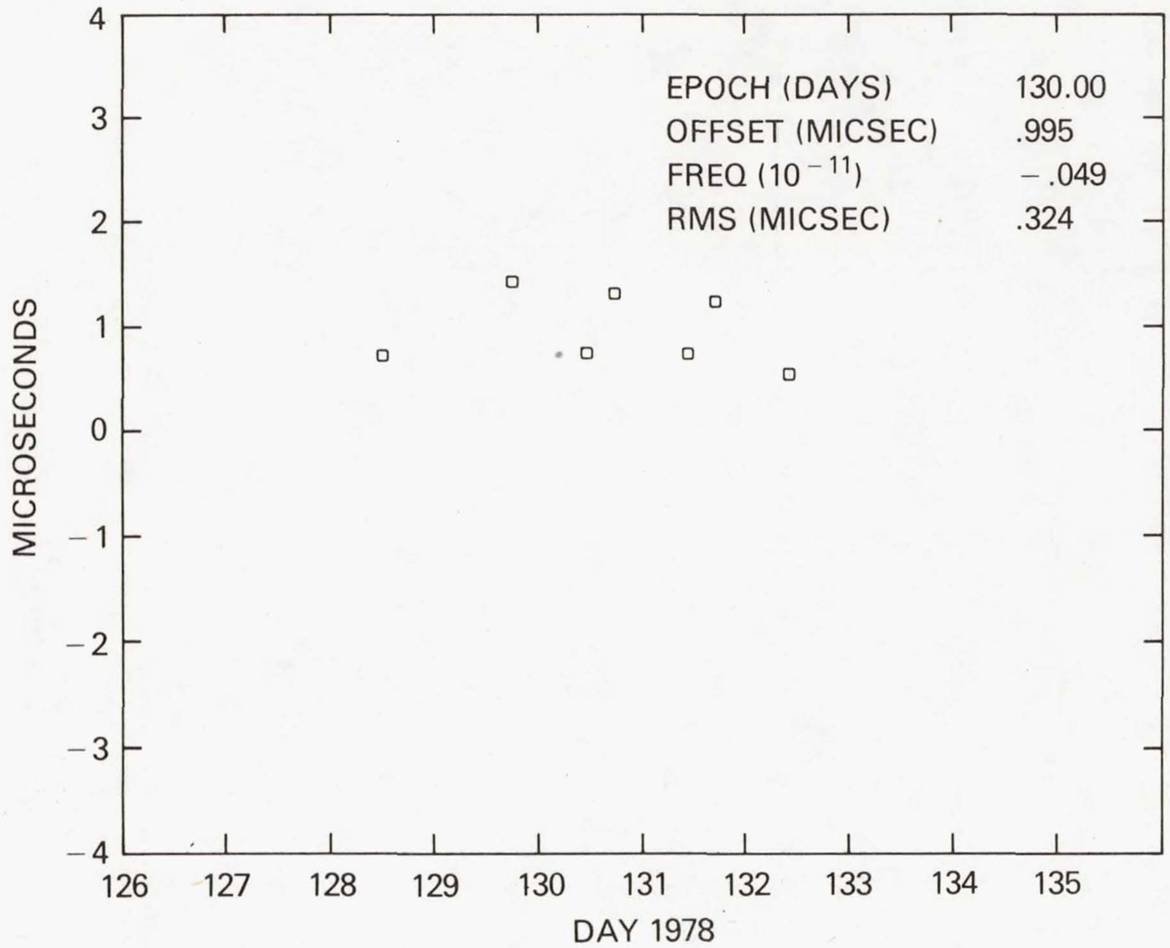


Figure 5 Time Transfer Results from Cerga, France
(USNO,MC1)-(CERGA)

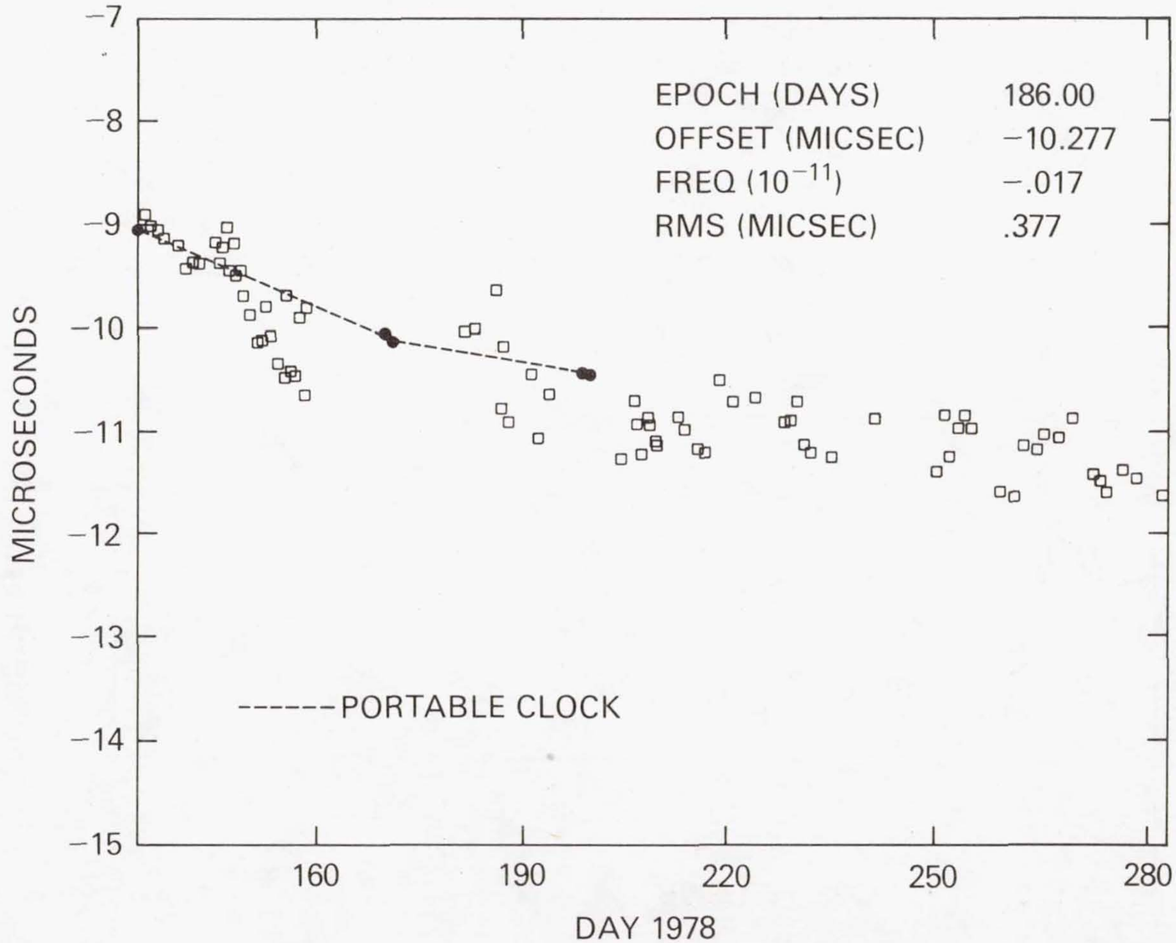


Figure 6 Time Transfer Results from the Institute for Applied Geodesy, Wettzell, West Germany
(USNO,MC1)-(IFAG)

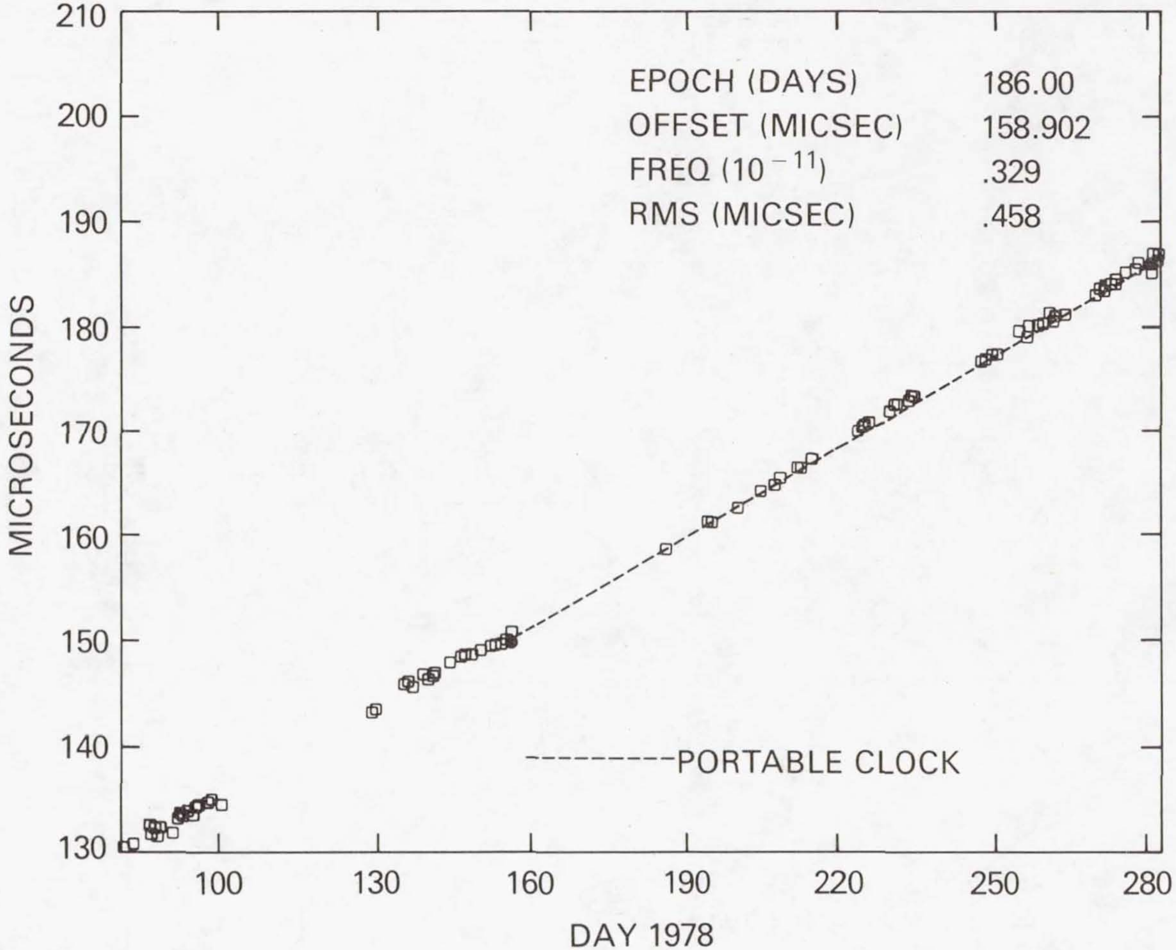


Figure 7 Time Transfer Results from the Division of National Mapping, Australia
(USNO,MC1)-(AUS,DNM)

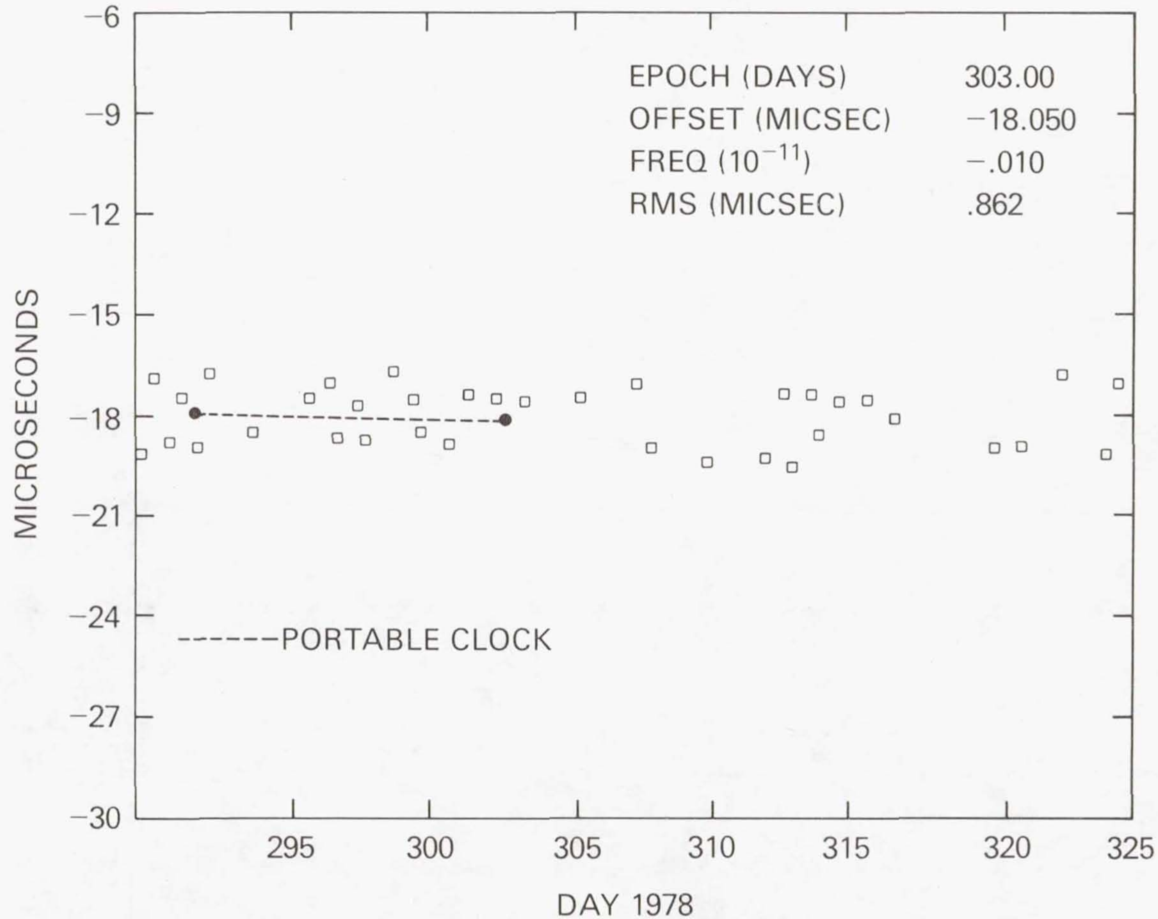


Figure 8 Time Transfer Results from RRL, Japan (USNO,MC1)-(RRL)

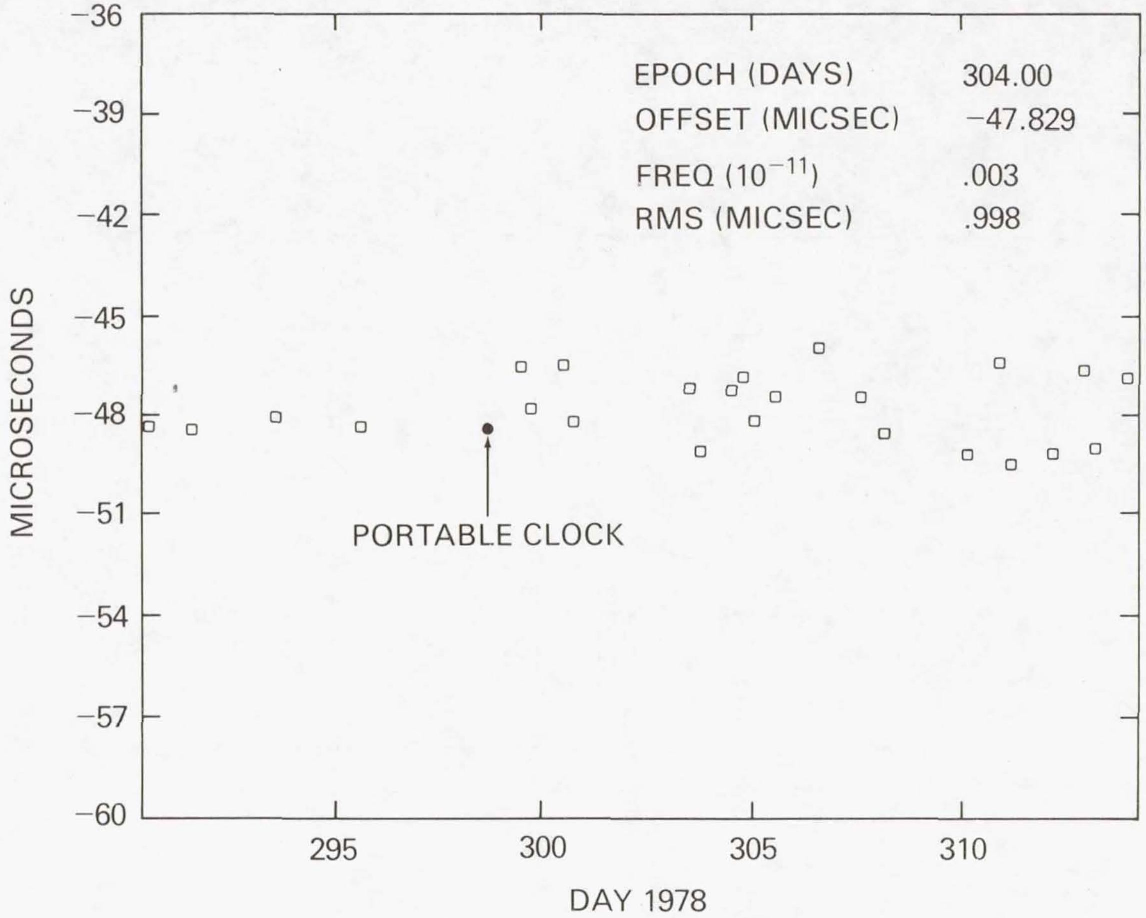


Figure 9 Time Transfer Results from NRLM, Japan
(USNO,MC1)-(NRLM)

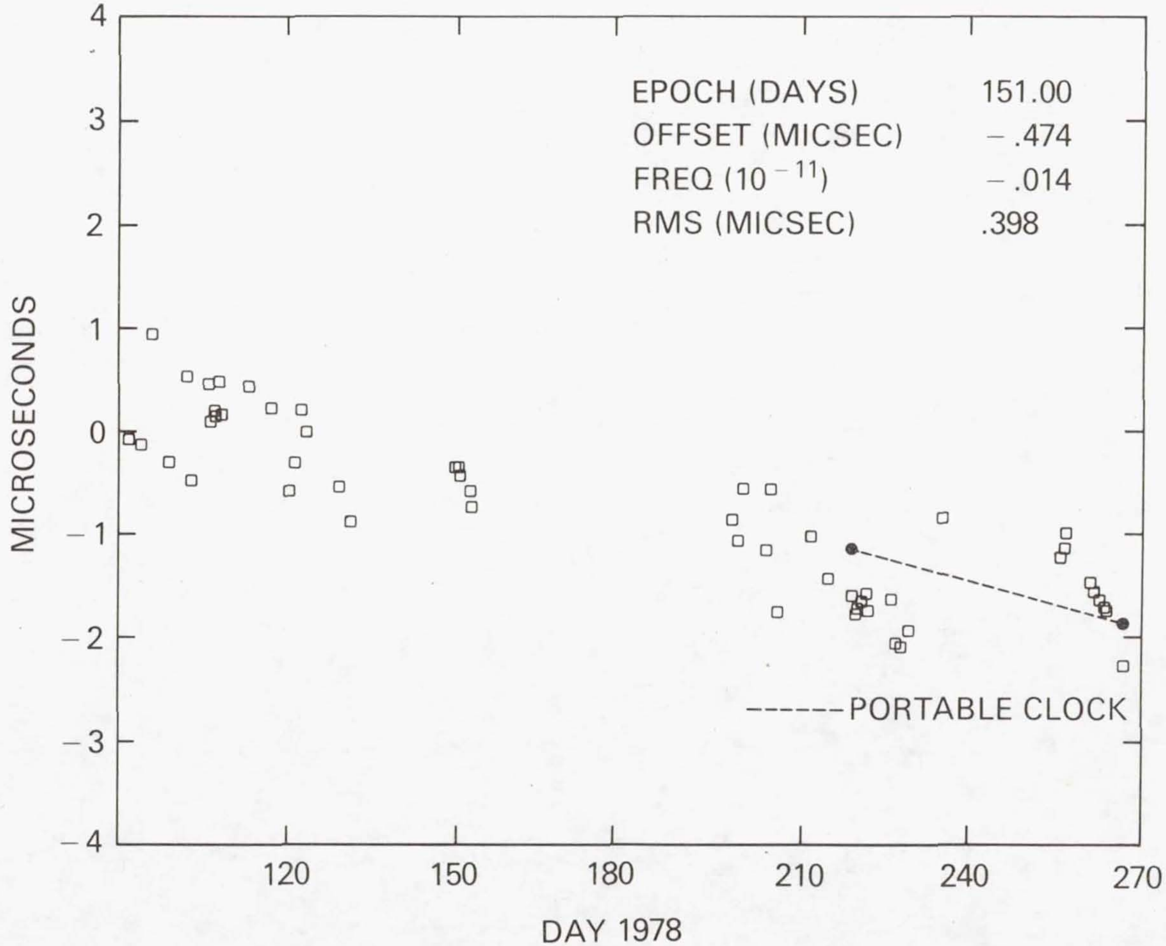


Figure 10 Time Transfer Results from NBS in Colorado, U. S.
(USNO,MC1)-(NBS)

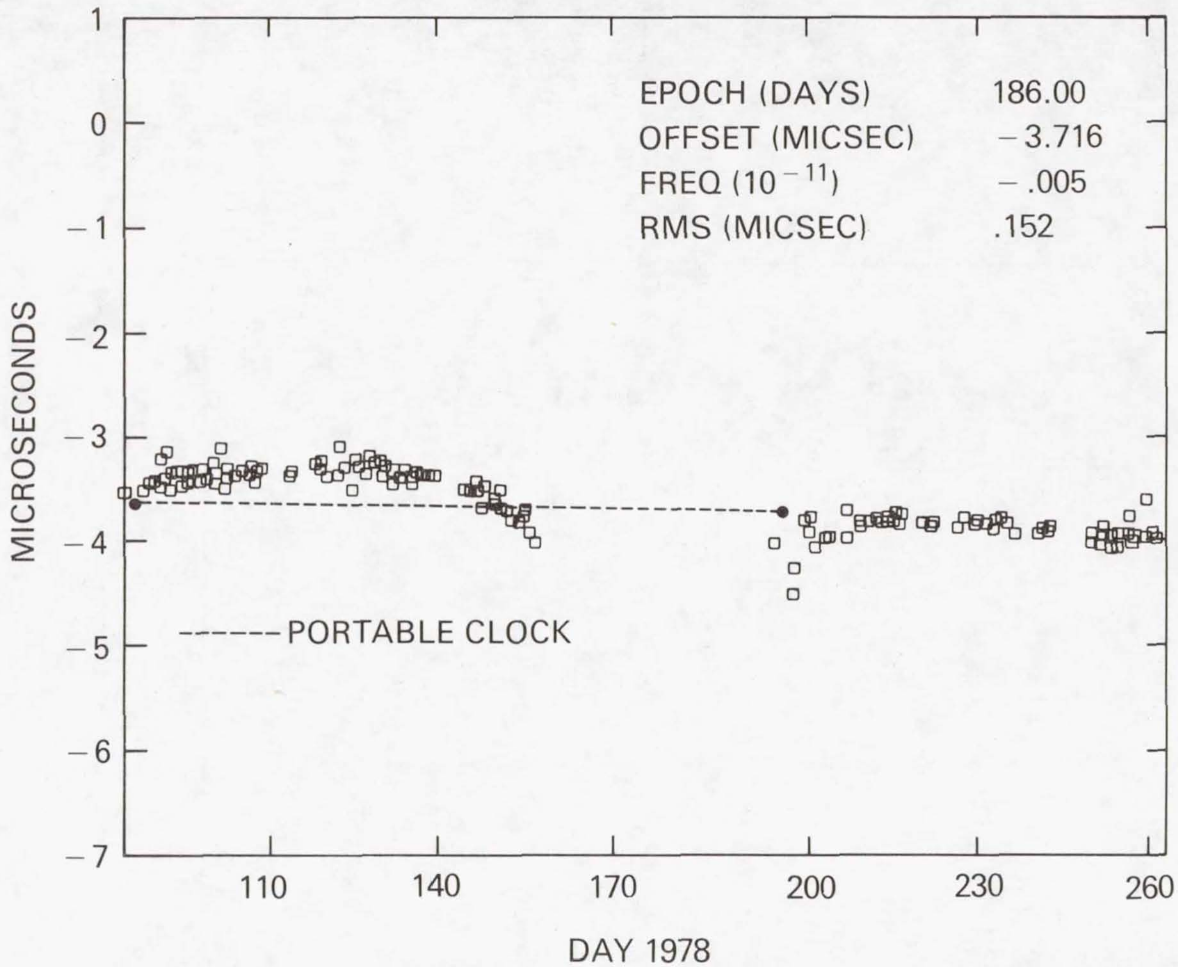


Figure 11 Time Transfer Results from National Research Council, Canada

(USNO,MC1)-(NRC)

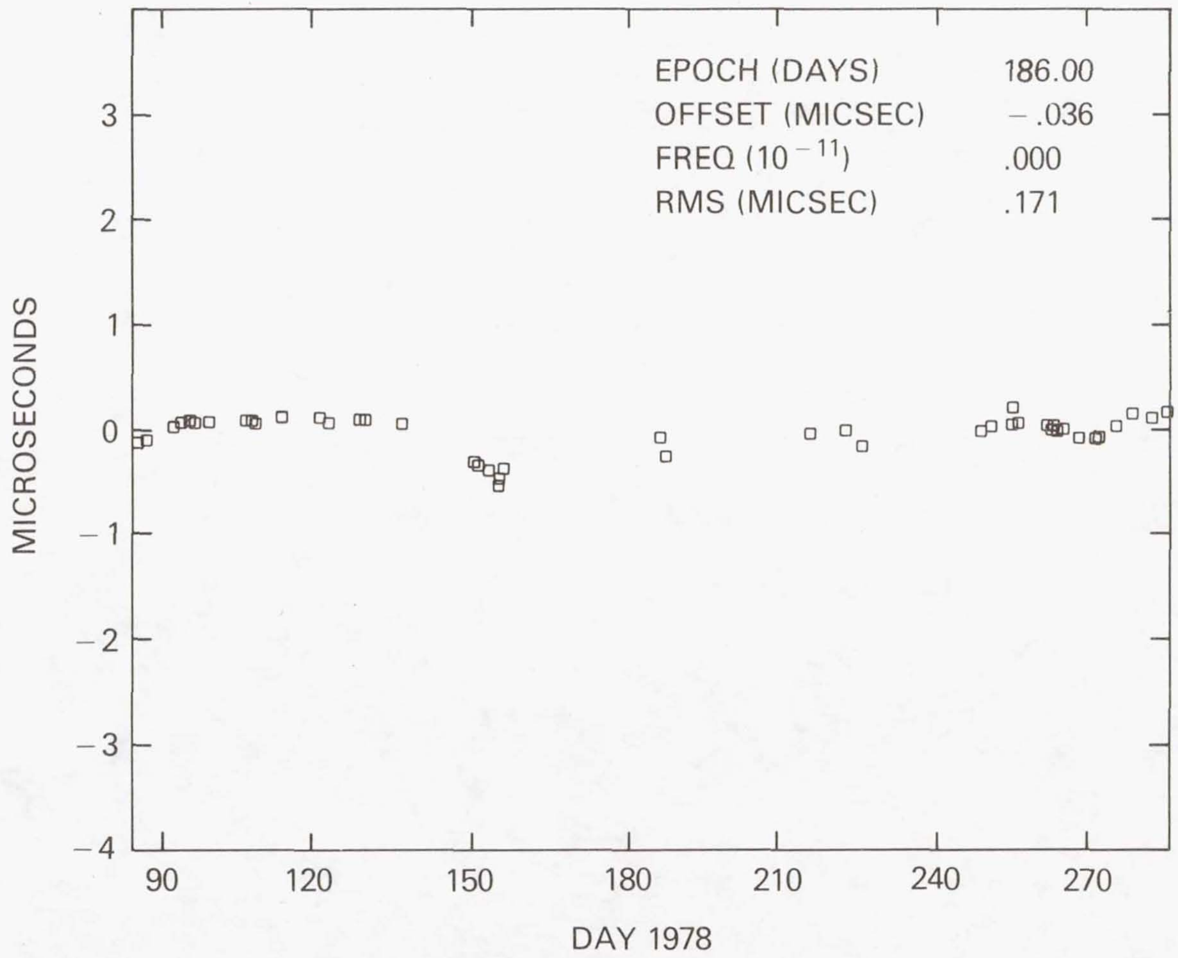


Figure 12 Time Transfer Results from USNO
(USNO,MC1)-(USNO,MC1,NTS)

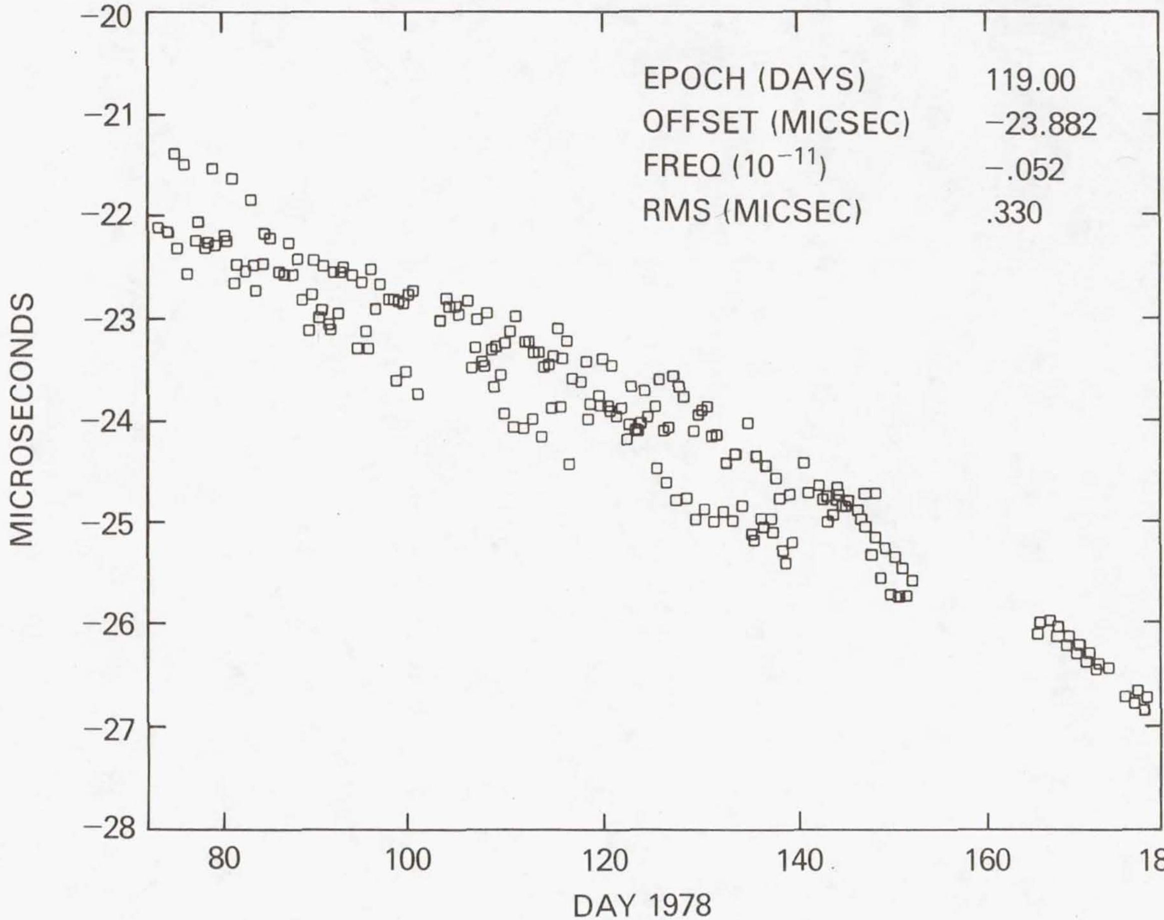


Figure 13 Time Transfer Results from Panama
(USNO,MC1)-(PMA)

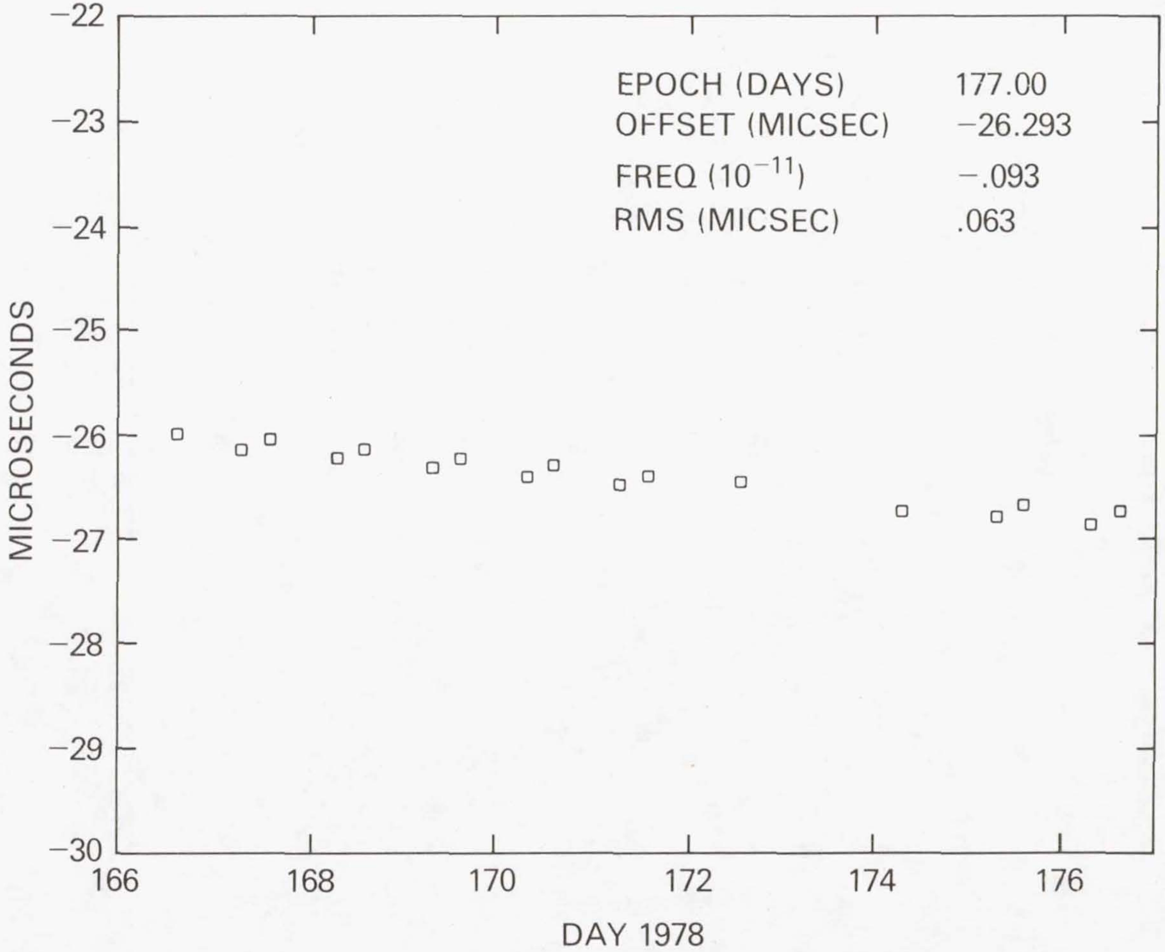


Figure 14 Panama Results (12 day segment)

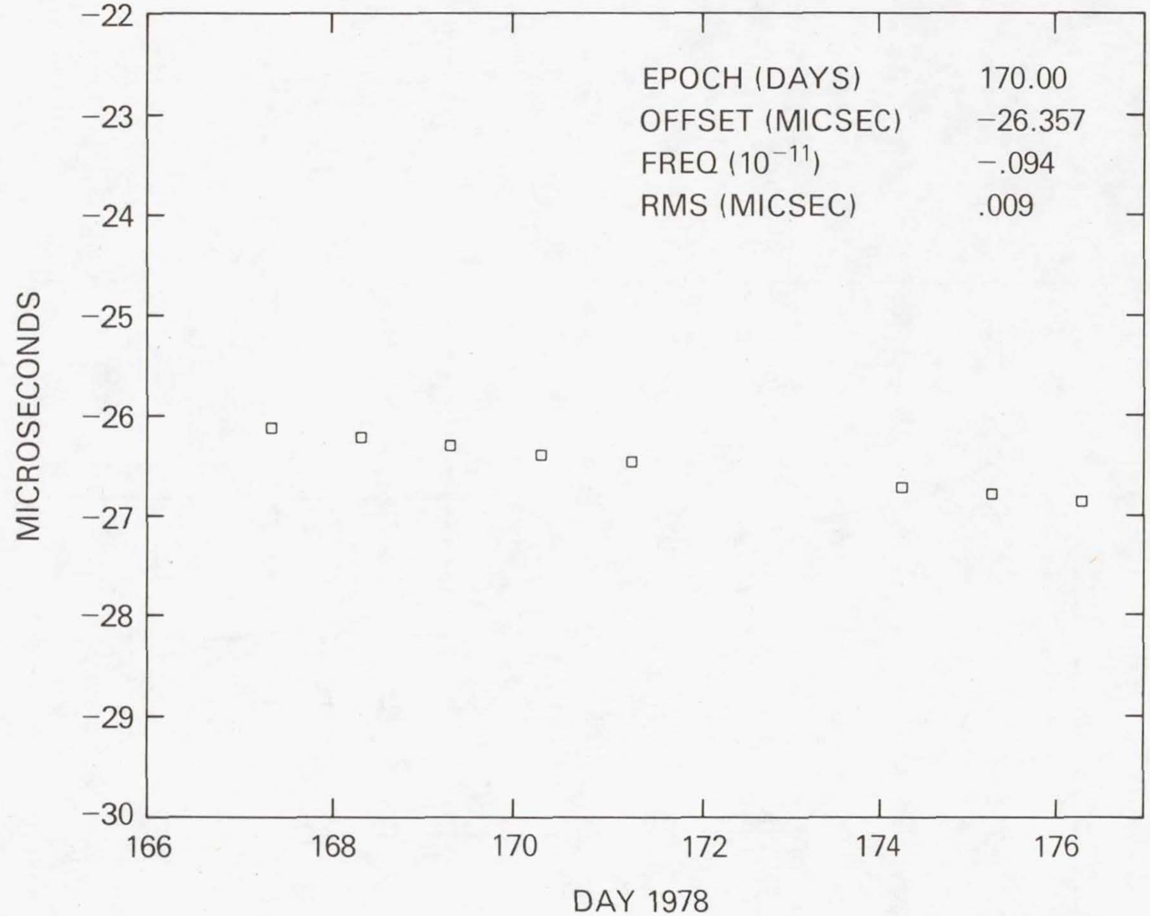


Figure 15 Panama Results (same successive revolution)

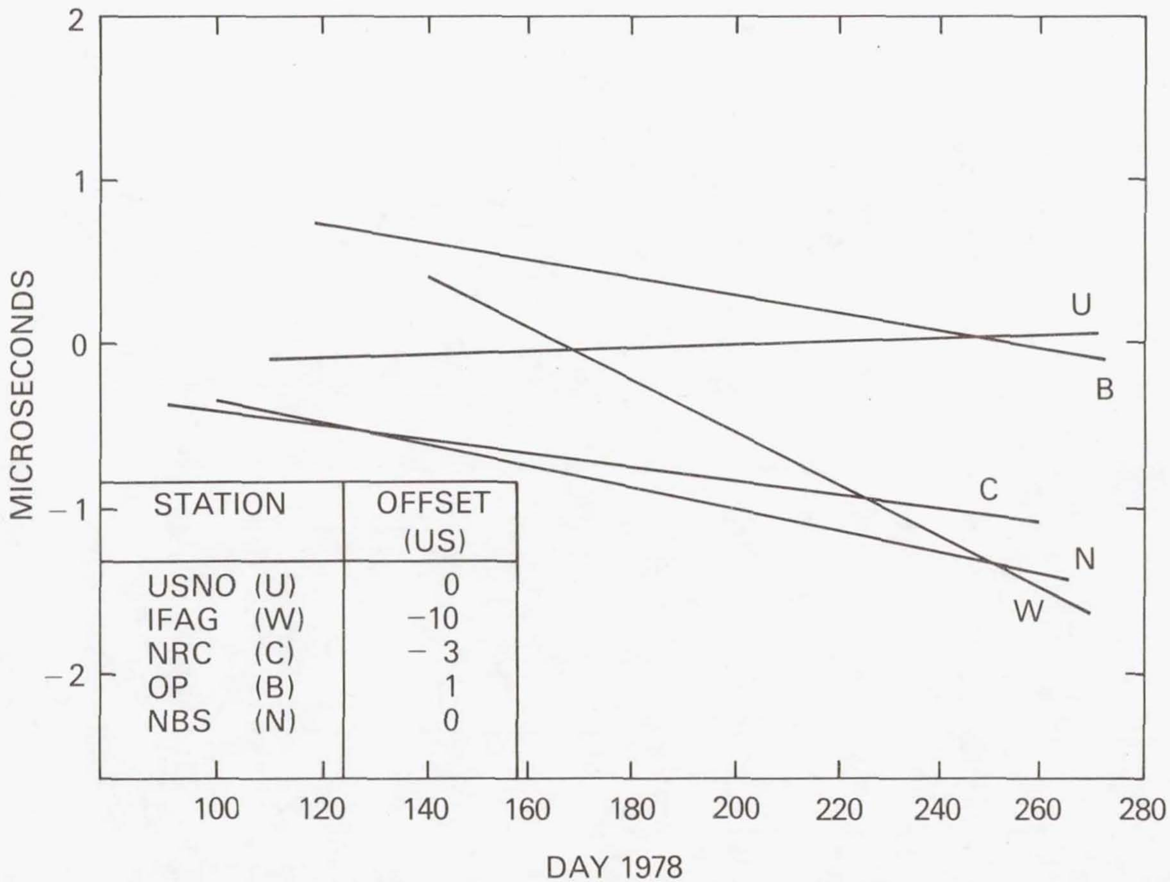


Figure 16 Time Comparisons via NTS

Questions and Answers

DR. DAVE CURKENDALL, Jet Propulsion Lab:

The data actually looks quite a bit better than the title of "Sub-microsecond." Would you care to venture just how good you think it is, really? How good do you think you have achieved time sync?

MR. MCCASKILL:

I would really prefer not to give an answer to that except that we are very well encouraged with the results. I don't really want to give an absolute number on it. In terms of what the GPS system might do, that slide that showed the nine nanoseconds, provided we can get all the biases out, which are fairly small and fairly well known right now, then probably in the neighborhood of 10 to 20 nanoseconds with no problem. Yes, sir.

DR. GERNOT M. R. WINKLER, U.S. Naval Observatory:

Since you don't want to say anything I would just like to give you my impressions on the basis of your data.

It appears that the NTS-1 comparisons have a noise level of about 150 nanoseconds, just because of the time transfer to the satellite and down. That is derived from your data of that closed loop, with the NTS receiver at the observatory.

The reason why the more distant stations show consistently larger sigmas, the larger the distances, is simply because NTS-1 only has a relatively poor frequency standard, and you depend on increasingly longer extrapolation times for the greater distances. So therefore, the data with Australia and with Japan are larger than one half microsecond, while those which are close by, such as NRC, have been down to 180, 200 nanoseconds. That is the NTS-1 results.

However, NTS-2 seems to be better by at least an order of magnitude because you have a cesium standard and therefore are much less sensitive to long extrapolation times. Also, it appears to me that you have a much better signal to noise; and I wonder whether you have any comments to that.

MR. MCCASKILL:

The only comment I would like to add to that is that the measurements with NTS-1 were made using a single frequency at 335 megahertz.

In those measurements we presented, there is no correction for the ionospheric delay, which is quite considerable. And in the NTS-2 results, we did have dual-frequency measurements available.

I certainly appreciate the comments concerning the quartz oscillator, but I think most of it is due to the ionosphere. Even though

quartz is certainly not as good as cesium, I believe the ionosphere is the major contribution there.

DR. TOM CLARK, NASA Goddard Space Flight Center:

I will take the prerogative of making a couple of instant analysis comments also. I notice a great deal of odd-even effect in the data points taken on the orbits, one side and the other, which to me says either a satellite ephemeris error and/or station location errors are contributing considerably to the scatter you are seeing on the plots.

Second of all, another one which I observed as it went by very quickly, the comparison of NRC to CBD versus the comparison of USNO to CBD both. The first half of the plot showed a very systematic parabola which looked just the same on the two of them, which indicates that the common clock, the CBD clock, was the one which was setting that part of the curve. And in fact, you might find it more instructive to subtract out CBD, and do a direct USNO-NRC comparison, which could then be tied into the previous paper. Do we have some other comments?

DR. BILL RECKERT, Rockwell:

I would like to know if there is any data that has come in on spacecraft now-- You said something about some intermittent problems. Is there another frequency that is being transmitted down in regard to useful data?

MR. MCCASKILL:

The answer to your first question is there is no data being acquired right now. It was on briefly last week. The last time it was on was during the eclipse period, which happens every six months; and it is not fully understood why it is on then not the rest of the time, but the next one comes up in about February. So we are hoping to see some more data from it, but there is no way to promise data.

DR. RECKERT:

What frequency is that data being transmitted on?

MR. MCCASKILL:

The NTS-2 results you saw were two frequencies at 335 and 1580 MHz, which were of course combined to make the ionospheric correction.

DR. VICTOR REINHARDT, NASA Goddard Space Flight Center:

Were both the NTS-1 and the NTS-2 data taken with the same type of receiver?

MR. MCCASKILL:

Some of the data were taken with the same type of receiver. However, part of the NTS-2 data were taken with a different type of receiver. Two different receivers were used.

DR. REINHARDT:

Is that the nine nanosecond data, two receivers?

MR. MCCASKILL:

Sir? Which--

DR. REINHARDT:

On the nine nanosecond data, two receivers were used?

MR. MCCASKILL:

That is correct; the other type.

DR. REINHARDT:

Thank you.

DR. WILLIAM KLEPCZYNSKI, U.S. Naval Observatory:

I am not sure I understand something which was said earlier in conjunction with the slide I saw. Earlier, Dr. Costain indicated that through the Hermes data, he came to the conclusion, or verified the conclusion, that NBS and USNO are going at the same rate, and NRC is running at a little bit different rate in the other direction. But your slide indicates that both NBS and NRC are running off with respect to the Observatory. I am wondering whether we are comparing the same thing: Are we comparing time scales, or individual cesiums, or what--I am not sure I understand.

MR. MCCASKILL:

Let me answer the question in two parts. First, we have not had a chance to cross-check with the Hermes results. The basic measurement is one of a start-minus-stop measurement; that is, as if you started with USNO Master Clock Number One and stopped with the result. So when you see a slope, it is a slope of the start-minus-stop measurement.

DR. WINKLER:

I think that both of you have indicated that there is no frequency difference between NBS and USNO, but that there is a frequency difference between NRC and both of us. Only that value is different; it is something like 1.1 in 10^{13} for the Hermes results, and only five parts in 10^{14} for your NTS results. This is my recollection. So they are in the same direction, and I don't think there is this discrepancy as Dr. Klepczynski seems to have seen.

You can go to the same slide, and there was no frequency difference between NBS, and there shouldn't be.

DR. CLARK:

Since it is ordained that there is no frequency difference between Boulder and Washington, let's go on.

Page intentionally left blank

SUBMICROSECOND COMPARISON OF INTERCONTINENTAL CLOCK SYNCHRONIZATION BY VLBI AND THE NTS SATELLITE*

William J. Hurd
Jet Propulsion Laboratory
California Institute of Technology

Jim Bussion
Jay Oaks
Thomas McCaskill
Naval Research Laboratory

Schuyler C. Wardrip
NASA Goddard Space Flight Center

Hugh Warren
Bendix Field Engineering Corp.

Gary Whitworth
Applied Physics Laboratory

ABSTRACT

The intercontinental clock synchronization capabilities of Very Long Baseline Interferometry (VLBI) and the Navigation Technology Satellite (NTS) were compared in May 1978 by using both methods to synchronize the Cesium clocks at the NASA Deep Space Net complexes at Madrid, Spain and Goldstone, California. The VLBI experiments used the Wideband VLBI Data Acquisition System developed at the NASA Jet Propulsion Laboratory. The NTS Satellites which were designed and built by the Naval Research Laboratory were used with NTS Timing Receivers developed by the Goddard Space Flight Center. The two methods agreed at about the one-half microsecond level. The VLBI system also obtained long term stability information on the HP5061A-004 Cesium standards by measuring $\Delta T/T$ over four 3-4 day intervals obtaining stability estimates of $(1 \pm 1) \times 10^{-13}$ for the combined timing systems.

*This paper presents the results of one phase of research carried out (in part) at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

I. INTRODUCTION

A series of experiments were conducted in May 1978 to compare the intercontinental clock synchronization capabilities of the Navigation Technology Satellite (NTS) time transfer system and a Very Long Baseline Interferometry (VLBI) system which is in use in the NASA Deep Space Net. The purpose of the experiments was to provide independent verification of the accuracy of both systems. This verification was accomplished at the one-half microsecond level.

The experiments were conducted between the 64m Deep Space Stations at Goldstone, California (DSS 14) and Madrid, Spain (DSS 63). The VLBI experiments used the Wideband Digital VLBI Data Acquisition System (WBDAS) developed at the NASA Jet Propulsion Laboratory, which has been in routine use in its present configuration since January 1978. The satellite time transfer experiments used the NTS-1 satellite, designed and built by the Naval Research Laboratory, and the NTS Timing Receivers developed by the Goddard Space Flight Center. The NTS receivers were brought to the DSS's specially for these experiments. Unfortunately the NTS-2 satellite was not available for the experiments, as use of this satellite might have improved the accuracy of the intercomparison by an order of magnitude.

II. Experiment Configuration

The configuration of the VLBI data acquisition system and the NTS time transfer receiver at each DSS is shown in Figure 1. Of particular interest is the clock system. The primary frequency standard at each station was a HP5061A-004 Cesium oscillator. Various frequencies are derived from the reference in a Coherent Reference Generator, including 50 MHz which is used in the VLBI system, and various coherent timing signals are made available at the output of the Time and Frequency Assembly. For the purpose of this experiment, the station reference clock is the 1 pps signal at the output of the TFA, since both the VLBI and the NTS systems connect to the station timing system at this point. In comparing the results for the two systems we accounted for the cable delays according to the specified or measured physical lengths of the cables, and we measured the delay from the 1 pps input to the 1 pps output of the NTS receivers, which is the reference for these receivers. The delay within the WBDAS units is less than 20 ns and is not significant.

III. NTS Time Transfer Results

Time transfer using a NTS satellite is accomplished by using the stable oscillator on board the satellite as a portable clock, and reading this clock over a microwave link as the satellite passes near the various ground stations. The two major fundamental sources of error in the time transfer are the instability of the oscillator on the satellite, and propogation delays in the ionosphere. Both of these error sources affect the time transfer measurement both directly, and indirectly through errors in the satellite orbit. Raymond, et al, [Ref. 1] describe the timing receiver used here, and present the results of some time transfer experiments between Rosman, NC, and the Naval Research Laboratory, using the NTS-1 satellite which was also used here. These experiments demonstrated rms errors of 86 ns with respect to a portable clock. Errors over intercontinental distances are somewhat more due to time separation, orbit errors, and larger ionospheric effects. The ionospheric effect and its uncertainty is often on the order of 1 μs at the radio frequency of 335 MHz, but this error source tends to cancel when the time and space separation are small.

The NTS time transfer measurements reported on here were more or less adjunct to the six nation cooperative experiment described by Buisson, et al, [Ref. 2]. The receiver used at the Madrid station was the same as the one used at Bureau International de l'Heure, France, and a spare receiver was used at Goldstone, California. As indicated in Figure 1, the receivers were installed in the DSS control rooms, with the NTS antennas on the roofs. The positions of the antennas were measured to within a few feet with respect to benchmarks at the stations, and errors in the antenna coordinates are not expected to contribute significantly to errors in the results. The data were processed in the same manner as in the six nation experiment, thereby estimating the offsets in the DSS clocks with respect to the USNO master clock C8D, at the Naval Observatory.

Figure 2 shows the results of thirteen measurements made at Goldstone from day 145 (May 25) through day 151 (May 31) of 1978. The results are corrected for a delay of 0.232 μs from the Goldstone clock reference point to the NTS receiver clock. A least squares linear fit to the data results in an offset USNO minus Goldstone of -0.688 μs at 0 hours on day 147, with a rate offset of -0.9×10^{-12} and an rms residual of 0.341 μs.

Ten measurements were conducted at Madrid during the same time frame, with the results shown in Figure 3. These results are corrected for a delay of 0.279 μs from the DSS clock reference point to the NTS receiver output. The least squares fit indicates an offset USNO minus Madrid of 8.593 μs at 0 hrs. on day 147, with a rate offset of -0.28×10^{-12} and rms residuals of 0.226 μs.

IV. VLBI System Description

Station clock offset is one of the many parameters which can be estimated by Very Long Baseline Interferometry. The random radio signal from an extragalactic radio star is observed at two antenna stations. Because the antennas are widely separated and the Earth is rotating, there is a time varying time delay between the arrival of the signal at the two stations. This time delay and its derivative can be estimated from the geometry, and can be measured by cross-correlating the signals received at the two stations. Because the arrival of the signal is time-tagged by the clocks at the stations, the difference between the measured and the predicted time delays forms an estimate of the offset between the station clocks.

The Wideband VLBI Data Acquisition System (WBDAS) has been described elsewhere (Ref. 3), together with results of experiments held in 1976, so we present only a brief description here. A simplified block diagram of the WBDAS is given in Figure 4. As shown in Figure 1, the system interfaces to the standard DSN receiving system at the 55 MHz $\pm$ 18 MHz output of the Block IV receiver. The receiver output is digitally demodulated to baseband by sampling at 50 MHz in each of two phase-quadrature 3-bit analog-to-digital converters. The A/D converter outputs are then low pass filtered, if desired, by summing N consecutive samples in a digital integrate-and-dump filter. These experiments used both unfiltered sampling, and filtering with $N = 3$, for a filter bandwidth of 16 2/3 MHz, which was a reasonably good match to the receiver system bandwidths.

The digital filter outputs are quantized to 1-bit and stored in a high speed buffer of 4096 bits. When the buffer is full, which takes about 120 μs for $N = 3$, sampling is inhibited and the buffer is emptied through the control computer onto digital magnetic tape. The total data rate onto magnetic tape is 57 kb/s, consisting of 14 bursts of 4096 bits. The control computer utilizes knowledge of the radio source position to predict the geometric signal delay from the source, and controls

the hardware buffer so that the same segments of the signal wavefront are sampled and recorded at both stations.

The utilized receiver bandwidth of 16 2/3 MHz is sufficient to achieve measurement resolutions of under 10 ns for any radio source which is strong enough to be detected. Resolution of about 1 ns is achieved with strong sources, using 1 minute of data (3×10^6 bits).

The accuracy of the system is limited primarily by propogation uncertainty in the ionosphere, which is often 20 ns at the S-band receiving frequency of 2290 MHz; by uncertainty in the Earth's orientation (UT1), which causes errors of about 5 ns; by errors in the positions of the radio sources; and by receiving system delays. We currently estimate the total day to day consistency in results to be about 30 ns; and the constant bias due to unknown but constant receiving system delays to be another 40 ns, for an estimated total error of 50 ns. (It is possible that the error in the receiving system delays is greater than 40 ns, because the delays have not been measured, but were estimated from cable length specifications).

V. VLBI Results and Comparison to NTS Results

Four VLBI clock sync experiments were conducted on May 15, 20, 24, and 27, 1978. The last three experiments consisted of from 8 to 13 total observations of 7-11 radio sources over total time spans of 1.5-3 hours. On May 15, due to operational problems, only two sources were observed, about 10 minutes apart. Despite the discrepancy in the amount of data, the expected clock sync errors are about the same on all days, except that the expected error on May 15 is slightly larger. As shown in Figure 4, the computer associated with the WBDAS is interfaced to data transmission lines. We used this capability to transmit some of the data from Madrid to JPL, and we processed this data between experiments to provide confirmation that the stations were properly configured.

Figure 5 shows the final clock offset estimates for the four days. The results are compensated for all known clock and signal delays and are expressed as Goldstone clock minus Madrid clock. A linear least squares fit to the data yields an estimated clock offset of 8.775 μs at 0 hrs. on day 147, with a rate offset of 0.33×10^{-12}. The rms of the residuals is 20.7 ns, and the sample standard deviation is 29.3 ns, with the difference due to estimating two parameters with only four data points. This

is compatible with our a-priori estimate of day-to-day consistency of 30 ns.

Also shown in Figure 5 is the NTS time transfer experiment estimate of the clock and clock rate offset between the two stations. This estimate is 9.281 μs at 0 hrs. on day 147, with a rate offset of 0.62×10^{-12}, which is just the USNO-Madrid result of Figure 3, minus the USNO-Goldstone result of Figure 4. The difference between the VLBI and the NTS estimates is 0.506 μs at 0 hrs. on day 147, with a rate offset of 0.29×10^{-12}. Day 147 was chosen as the reference epoch because both experiments were in progress at that time.

The difference of 0.5 μs between the two experiments is probably mainly due to the ionospheric effects on the NTS measurements, both directly and through errors in orbit determination. There may also be a larger constant error in the estimated station delays for the VLBI experiment than the anticipated ±40 ns. The difference between the rate estimate is within the error bounds of the NTS experiment.

VI. Oscillator Stability Estimate from the VLBI Data

The instabilities of the HP5061A-004 Cesium oscillators at the two stations can be bounded by using the VLBI results. The four experiments form three time intervals of 3 to 4.5 days. Differencing the clock offset estimates for successive experiments leads to frequency offset estimates of 1.85×10^{-13}, 4.26×10^{-13}, and 3.63×10^{-13}. Successive absolute differences between the offsets, divided by root 2, yield Allan variance σ's of 1.70×10^{-13} and 0.44×10^{-13}. The average of the two Allan variance pairs yields an average σ of 1.07×10^{-13} with a sample sigma of 0.9×10^{-13}.

We have estimated the combined instability of the two Cesium oscillators, the station time distribution systems, and the VLBI measurements, over 3-4 day intervals, to be approximately one part in 10^{13}, with an uncertainty of one part in 10^{13}. By increasing the time interval to about 10 days and by reducing the ionospheric effect on the VLBI measurements by use of X-band, long term frequency stability measurements at the 10^{-14} level seem feasible.

VII. Conclusion

By intercomparison of results, we have demonstrated the absolute accuracy of the NTS time transfer system and the WBDAS VLBI

system to be 0.5 μs or better between Goldstone, California, and Madrid, Spain. For the VLBI system, we have produced clock sync residuals which demonstrate day to day consistency at the 20-30 ns level, and the ability to use this system to measure long term frequency stability at the 10^{-13} level. Frequency stability measurements at the 10^{-14} level are indicated to be feasible.

REFERENCES

1. Raymond, L., et al, "Navigation Technology Satellite (NTS) Low Cost Timing Receiver," Goddard Space Flight Center, Greenbelt, Maryland. Technical Report X-814-77-205, August 1977.

2. Buisson, J., et al, "Submicrosecond Comparisons of Time Standards Via the Navigation Technology Satellites (NTS)," Proceedings of the Tenth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, US Naval Research Laboratory, Washington, D.C., November 28-30, 1978.

3. Hurd, W. J., "Preliminary Demonstration of Precision DSN Clock Synchronization by Radio Interferometry," in the Deep Space Network Progress Report 42-37, Jet Propulsion Laboratory, Pasadena, CA, February 15, 1977, pp. 57-68.

FIGURE 1. CONFIGURATION OF THE NTS RECEIVER AND THE VLBI SYSTEM IN A DSS

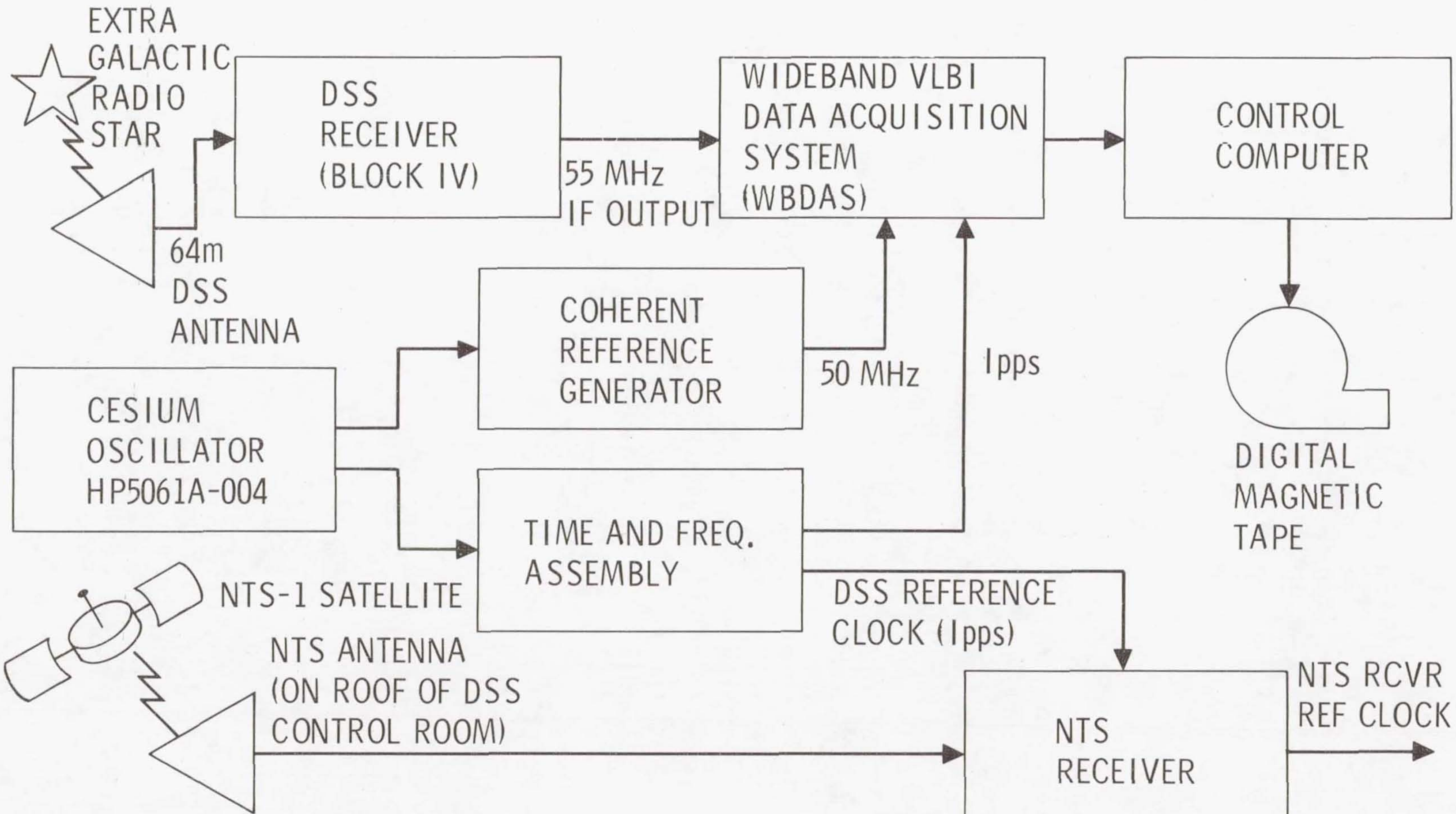


FIGURE 2. NTS1 TIME TRANSFER RESULTS FOR GOLDSTONE

USNOMC (C8D) - GOLDSTONE (DSS 14)

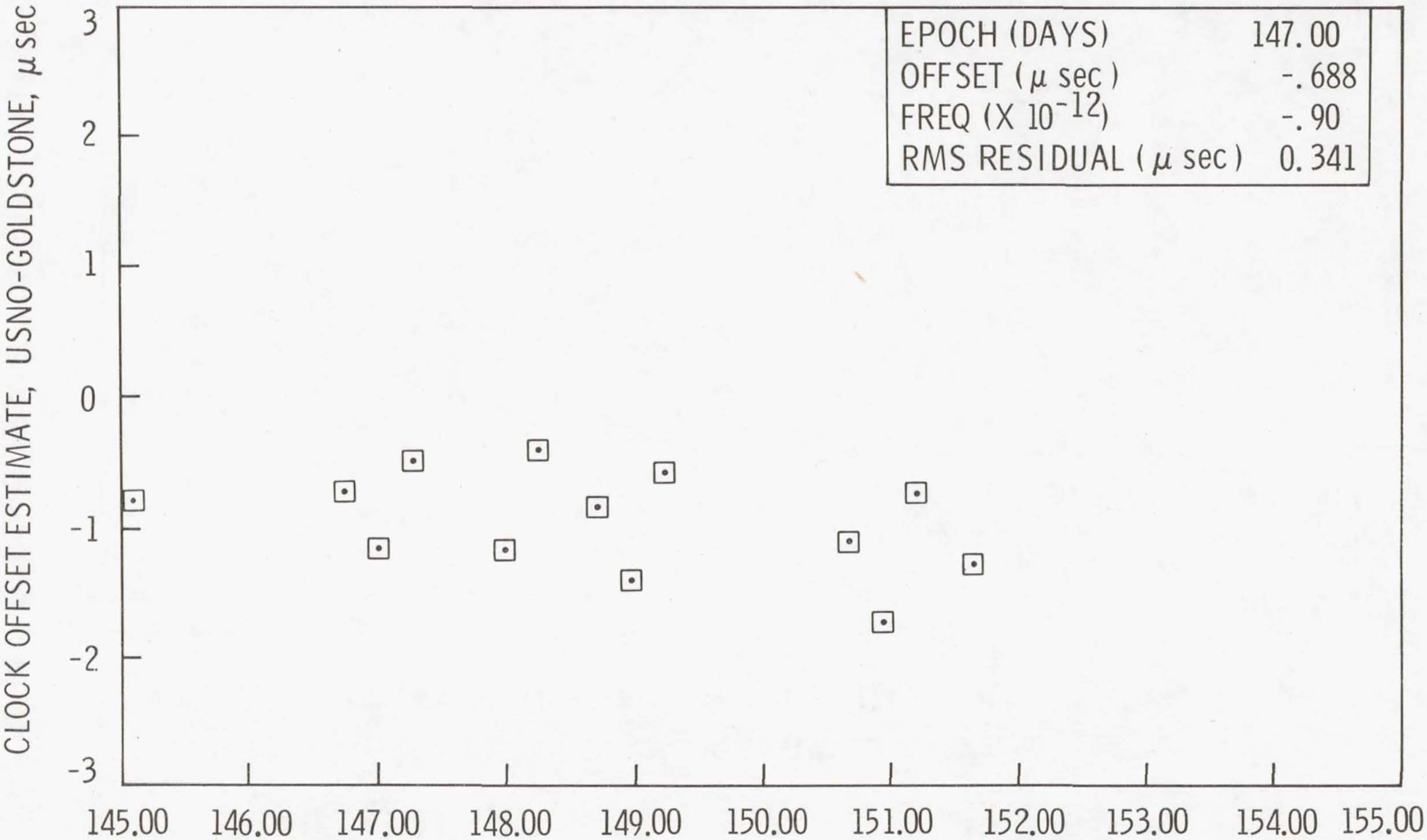


FIGURE 3. NTS 1 TIME TRANSFER RESULTS FOR MADRID

USNOMC (C8D) - MADRID (DSS 63)

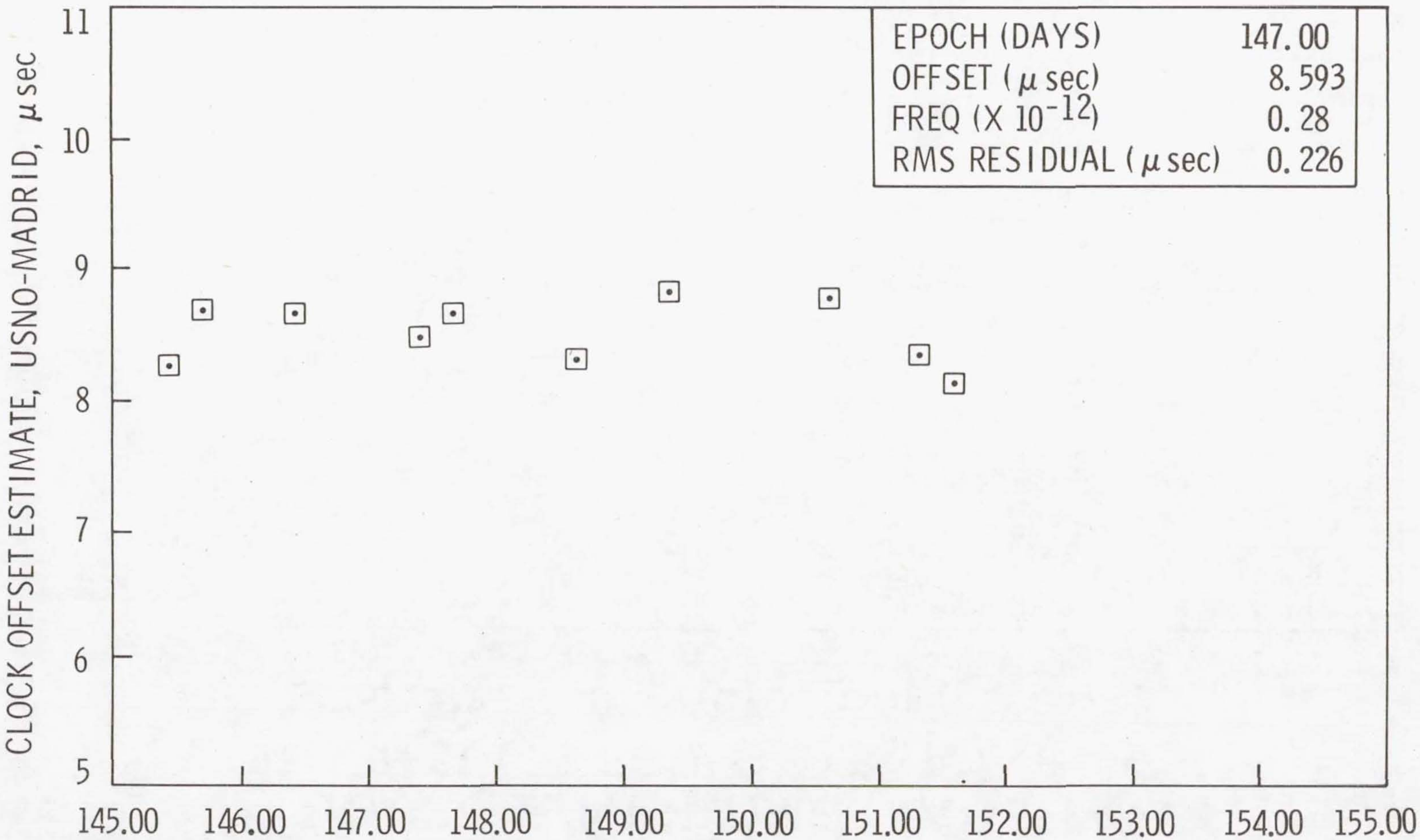
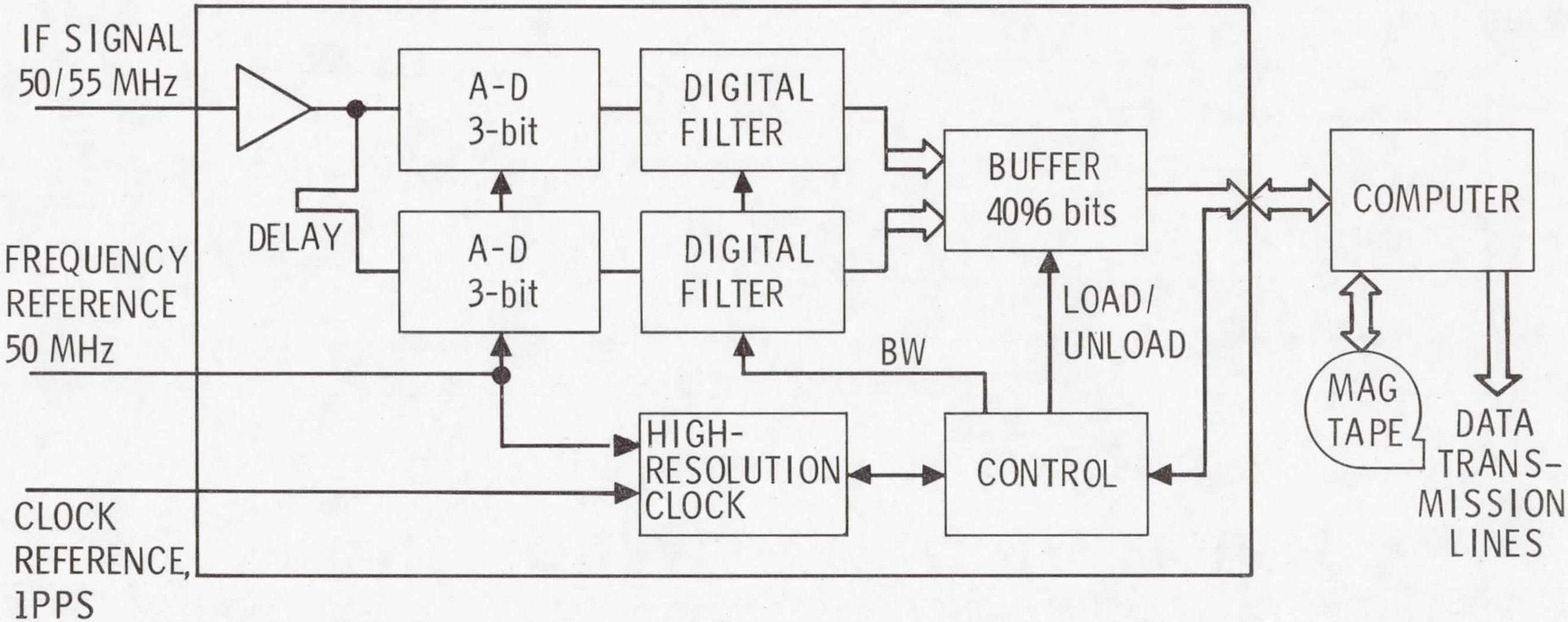


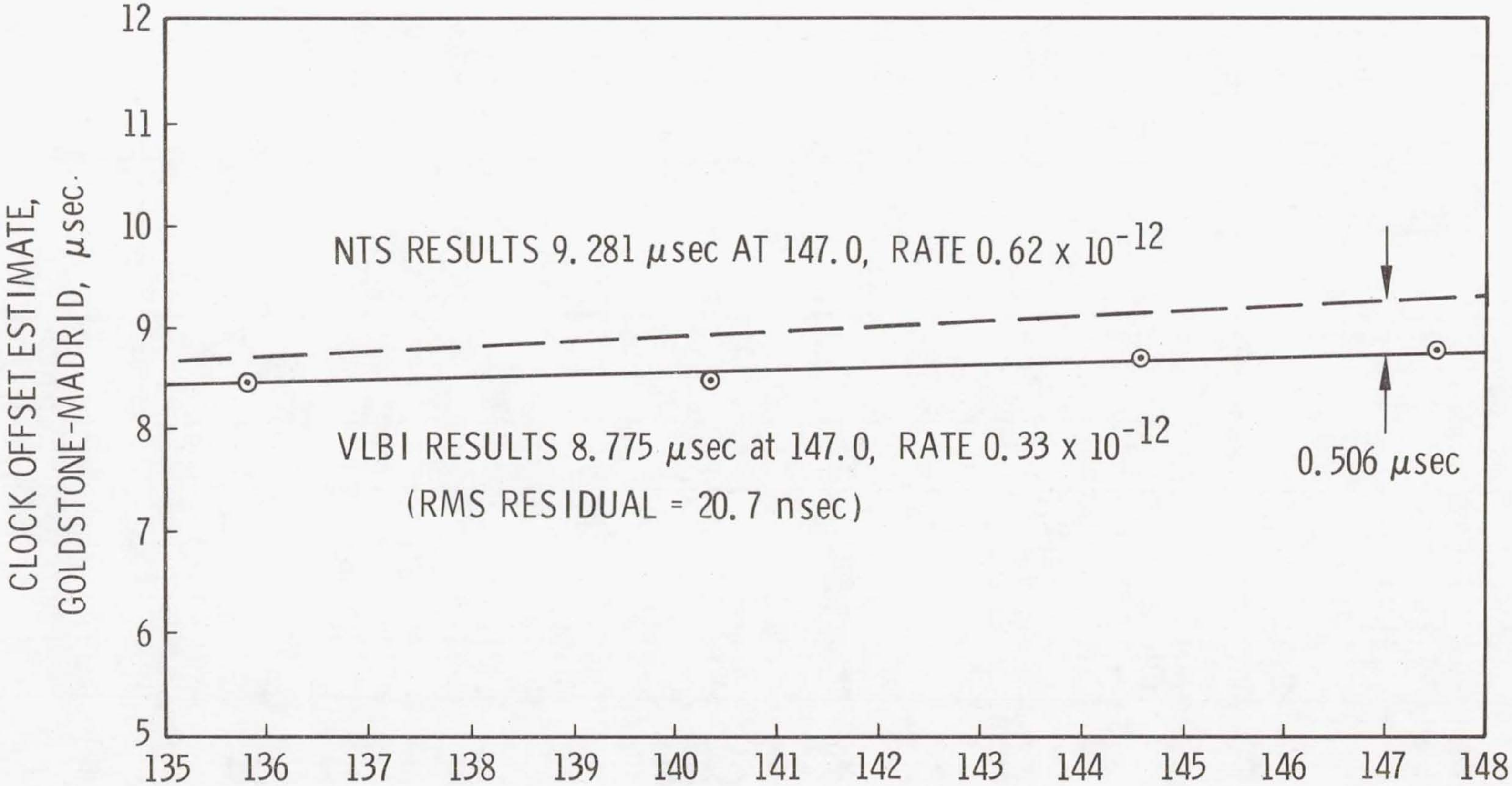
FIGURE 4. WIDEBAND DIGITAL DATA ACQUISITION SYSTEM BLOCK DIAGRAM



639

FIGURE 5. VLBI CLOCK OFFSET MEASUREMENTS, AND COMPARISON TO NTS RESULTS

GOLDSTONE (DSS 14) - MADRID (DSS 63)



Questions and Answers

DR. GERNOT M. R. WINKLER, U. S. Naval Observatory:

I think you are absolutely right. That comparison to one part in 10^{13} looks as good as you can possibly expect it to considering the variations of these clocks just on themselves and all your other circuitry.

But there is one very stupid question for which I hope you will forgive me, on account of the late state of the conference. I have heard so much about the absolute delay measurements from the antenna to the point where you make the timing. Why can't one have, I wouldn't say a portable clock calibration to the antenna tip, but why can't one inject at a reference point in the antenna a signal on a fixed and calibrated length, which is transported to that point, and then do the same thing at your clocking station? Could you give me an answer to that, please?

DR. HURD:

Well, Tom Clark is one of the world's experts on this, but from a JPL point of view, yes--we are implementing such a system. We have experimental copies of the MIT system, which we hope work, and they are supposed to be installed in the stations in the next six months or so, I believe. So by this time next year, we should have that calibration.

DR. CLARK:

Yes. At last year's PTTI meeting, I gave a paper which did use just such a calibrated system, which did calibrate all the cables through the system, and between Massachusetts and West Virginia. We reported on two 10-nanosecond clock synchronization experiments in conjunction with the Naval Observatory. So yes; it can be done, and has.

MR. LAUREN RUEGER, Johns Hopkins University, Applied Physics Lab:

One of the problems in doing what Dr. Winkler has suggested is, you must take into account the delays that the antenna itself introduces. In some of our instruments we have used, there was as much as 70 nanoseconds antenna delay.

DR. HURD:

Yes. There is an antenna multipath problem involved in here. One of the arguments for various VLBI systems is that the effect of multipath might be different on different systems. You have to inject your test signal as far forward in the system as you can.

One of the most unstable portions of the receiving system for VLBI is the traveling wave maser, the first amplifier. And we estimate that due to the way it is tuned that day, it may vary day-to-

day on the order of 10 or 20 nanoseconds, and we hope to be able to control this. And if this is the cause of errors in our measurements, then we have more motivation to force the stations to control this tuning, or calibrate it with the phase calibration sytem.

My system is really an attempt to make some useful measurements in the DSN without the need for the phase calibrator--up until the time when the phase calibrators are available, and in case they are not working someday.

CLOCK SYNCHRONISATION EXPERIMENT IN INDIA USING SYMPHONIE SATELLITE

Y.V. Somayajulu, B.S. Mathur, P. Banerjee, S.C. Garg, Lakha Singh, P.C. Sood and Tuhi Ram Tyagi, National Physical Laboratory, New Delhi-110012 (India)

and

C.L. Jain and K. Kumar
Space Applications Centre, Ahmedabad-380053 (India)

ABSTRACT

A recent clock synchronisation experiment between the National Physical Laboratory (NPL), New Delhi and Space Applications Centre (SAC), Ahemedabad, in India via geostationary satellite symphonie-II, stationed at 49^{o} E longitude, is reported in this paper. As only one satellite transponder was available for this experiment, the two way transmission of the clock pulses was carried out by switching the transmit-receiver roles at the two stations at 5 minute intervals to achieve a nearly simultaneous two way transmission. Taking into account all the additional delays, the results demonstrated a clock-synchronisation accuracy of better than 0.5 microseconds. A crystal based portable clock flown aboard an aircraft confirmed this clock-synchronisation within a microsecond.

INTRODUCTION

The feasibility of precise clock synchronisation on an inter-continental basis has been demonstrated and looks attractive for time transfer on a global scale (See Somayajulu, 1977, for a recent review). These experiments carried over the past decade and a half have increased the precision of clock synchronisation. Currently a two-way transmission using a microwave transponder aboard a geostationary satellite is considered to provide the greatest precision in synchronisation of two remote clocks.

With the availability of the French-German geostationary satellite Symphonie-II to India for telecommunication experiments under a bilateral agreement an opportunity is provided for initiating clock synchronisation experiments in India. The satellite was parked over the equator above 49° E longitude in June 1977 add is expected to be located

there for two years. Clock synchronisation experiments have been carried out from 17 to 30 April and 16th to 30 June, 1978 between National Physical Laboratory, New Delhi and Space Applications Centre, Ahmedabad via the Delhi and Ahmedabad earth stations. This report describes the details of these experiments and the results obtained.

DETAILS OF THE EXPERIMENTAL SET-UP

The symphonie satellite has two C-band transponders aboard of which one transponder was available for these experiments. The frequency for uplink transmission is 6347.5 MHz and is 4122.5 MHz for downlink reception. Both the Delhi (DES) and Ahmedabad (AES) earth stations are equipped with similar transmitting and receiving equipment. The relevant details are given in table-1.

Table 1. Details of Delhi andAhmedabad Earth Stations

	Delhi Earth Station	Ahmedabad Earth Station
Latitude (Geographic)	28° 36' 30"N	23° 1' 21.03"N
Longitude (Geographic)	77° 11' 0.0"E	72° 33' 52.48"
Height of the Antenna (above mean sea level)	240.7 m	74.2 m
Antenna Type & size	Parabolic dish Dia 10.7 metre	Parabolic dish Dia 14 metre
EIRP dbw (Max)	84.0	87.7
Transmit Freq.	6347.5 MHz	6347.5 MHz
Receive Freq.	4122.5 MHz	4122.5 MHz
IF Bandwidth	10 MHz	10 MHz
Modulation	FM	FM

A simplified block diagram of the experimental configuration at the two ground stations is shown in fig. 1. The C-band transponder on board the Symphonie satellite receives at a frequency of 6347.5 MHz and transmits at a frequency of 4122.5 MHz. The satellite transponder acts as a simple frequency translator without any signal processing. The bandwidth of the transponder is 90 MHz. The effective isotropic radiated power (EIRP) of the transponder is 29 dbw.

Each station in turn transmitted a continuous wave (CW) carrier that is deviated 240 KHz for a 30 μs period of time. The rise time of the clock pulse is 50 ns. This deviation is caused by a 30 μs clock pulse generated at a rate of 1 pps. The pulse shape is checked on an oscilloscope before starting the time measurement. The signal received from the remote station via the satellite is detected by the limiter/ discriminator in the receiving system at the ground station. The leading edge of the discriminator output pulse stops the time-interval counter (HP Model 5245 or 5248 L/M) which is started by the local clock pulse. In effect, the counter measures the elapsed time between the start pulse of the local clock and the stop pulse from the remote clock. The counter readings are recorded with 10 samples centred on each minute and averaged. The duration of the experiments each day was 90 minutes. Each station transmitted for 5 minutes while the other received; immediately after 5 minutes the transmit-receive roles are switched for the next 5 minutes. This was repeated every 1/2 an hr. so that a maximum of 6 sets of samples are obtained each day. The time difference between the two clocks located at the ground stations is determined from these readings.

SYNCHRONISATION OF DES ATOMIC CLOCK WITH MASTER CLOCK

The National Physical Laboratory (NPL), New Delhi, maintains the national time standard through a cesium clock HP Model 5061 A with option 004. NPL also transmits time signals derived from the standard clock on ATA-transmissions at carrier frequencies of 5, 10 and 15 MHz. Another cesium clock type HP 5061 A with option 004 was located at the Delhi Earth Station and was kept in synchronization with the master clock by the following methods:

i) Via tracking of ATA broadcast

From the ATA-time transmissions received at DES 1 PPs pulses are generated by zero-crossing detection technique. The delay between these pulses and the 1 pps pulses from the Cesium Clock located at DES is measured by the time interval counter; from this delay measurement and computing the

(ground wave) propagation delay between ATA (Greater Kailash, New Delhi) and DES, the DES cesium clock is adjusted for synchronism with the Master Clock within 50 μs.

(ii) Via Portable Clock

A 1 pps pulse is generated from the 100 KHz output of HP 105B type crystal oscillator. This 1 sec pulse is autosynchronised with the Master Cesium Clock. The synchronised crystal clock is immediately transported to DES where it is then used to synchronise the DES cesium clock to about 0.1 μs. The history of the crystal is well known before. At the end of each experimental period, the cesium clock at DES was transported to ATA for direct comparison with the master clock. The offset measured was consistent with that determinted by the portable crystal clock to within 0.1 μs.

Before the start of the experiments the AES atomic clock is synchronised to the Master Clock to within 1 ms using the ATA-transmissions received at Ahmedabad.

FACTORS AFFECTING THE PRECISION AND ACCURACY OF CLOCK SYNCHRONISATION

(a) Precision of Synchronisation

The precision of clock synchronisation, i.e., the minimum absolute time difference to which two Clocks could be synchronised, is determined essentially by the signal-to-noise power ratio at the receiver output which provides the stop pulse to the time interval counter and the counter error, the latter being taken as $\pm$ 1 digit. The finite signal-to-noise ratio causes a jitter of the leading edge of the clock pulse, thus causing an error in the measurement. The r.m.s. jitter in the arrival time of the clock pulse is given by

$$\delta t_{\text{rms jitter}} = \frac{tr}{(2^{S/N})^{1/2}} \qquad \ldots\ldots\ldots(1)$$

where tr is duration of the clock pulse and S/N is the signal-to-noise power ratio. In our case tr is 30 μs. The stated S/N ratio for both DES and AEA is 50 db min. Using these values, from equ. (1) we obtain an estimate of rms jitter as 67 ns.

The combination of the rms jitter and the inherent counter error will show up as random fluctuations of the counter output superposed on any systematic drift due to factors to be described in the next section. Thus the standard deviation of the counter reading is a measure of the precision of synchronisation.

(b) Accuracy of Clock Synchronisation

The accuracy of the clock offset measurement or synchronisation depends on the extent to which signal delays introduced in the total system are known or accounted for. The total signal delay time comprises the time delay introduced by the intervening electronic equipment, i.e., the transmitting equipment at the earth station, the satellite transponder, and the receiving equipment at the other ground station, the delay in the ionosphere and the troposphere, and the free space path delay between the ground station and the satellite. Thus contributing factors for the total signal-delay error budget are:

(i) Ionospheric and tropospheric delay errors

(ii) Satellite position error

(iii) Equipment delay errors

The signal delay is defined as the time required for an identifiable point in the signal waveform from entering the transmitting equipment to its reappearance at the output of the receiver. In the present case it is the 50% value of the leading edge of the 1 pps pulse.

i) Ionospheric and tropospheric errors

In the earth's ionosphere and the troposphere the signal delay is equal to the group delay i.e., the signal energy propagates with the appropriate group velocity. The group delay in the ionosphere is essentially proportional to the total electron content which has a diurnal variation, also a day-to-day variability and is a function of frequency. As the additional ionospheric delay is frequency dependent, an error in a two-way method will be caused by the inequality of the propagation time for two directions of the path. In high frequency approximation, the ionosphere delay 'd' is

$$d = \frac{40.3}{c f^2} \int_{Path} N_e \, ds = \frac{40.3}{c f^2} N \quad ---- \quad (2)$$

where

c : light velocity

f : frequency in Hz

Ne: election density per meter cube

N : total election content along the path

using eqns. (2), the error due to the ionospheric effect, E_I, for the up-link of 6 GHz and the down-link of 4GHz is given by

$$E_I = \frac{1}{2} (d_{DA} - d_{AD}) - 2.33 \times 10^{-27} (N_D - N_A) \quad ...(3)$$

where N_D , N_A are the total election contents (e^1/m^2) along the path from satellite to the DES & AES respectively and d_{DA} and d_{DA} are the delay from DES to satellite and to AES and from AES to satellite and DES respectively.

On the other hand the troposphere group delay is practically independent of frequency and is insignificant for elevation angles greater than 15° (In the present case the elevation angle is 45°). At the C-band microwave frequencies, treating the velocity propagation as essentially the velocity of light (2.9979×10^8 m/s), from earth station antenna to the satellite introduces an error of less than 5 ns in the group delay computation.

In the present experiment a two-way transmission is used. Although it is not strictly simultaneous, it is nearly simultaneous in the sense that the transmit-receive roles of the two stations are alternatevely switched after a 5-minute interval. During this time the ionospheric and tropospheric conditions are practically unchanged and therefore the propagation delay drops out in the final computation of the clock offset. The contribution due to any satellite motion is discussed in the next section. An explicit assumption involved is that the electromagnetic path between DES and AES is reciprocal. This assumption is not strictly valid because the uplink and downlink frequencies differ by 2 GHz. However, the differences in the path reciprocity are negligible at the C-band frequencies, amounting to less than 1ns.

ii) Satellite Position Errors

Symphonie satellite is not in an absolutely synchronous orbit. The motion of the satellite during the period of tests produces a steady and systematic change of the apparent time delay that is measured. The DES and AES transmit-receive roles are switched alternately for 5 minutes. During the period of measurements the satellite drift error is less than 1μs over the one minute interval for which time delay is measured. In any case, this systematic drift is taken into account by least square fit on a computer of the piecewise observations.

The error in the position of the ground stations does not exceed 1m and is hence inconsequential.

iii) Equipment Delay error

The time signals experience additional delay in passing through the electronic equipment in the transmit-receive chain, viz., the transmitting and receiving equipment at each ground station and the satellite transponder. In the present case the satellite transponder operating in the C-band uses very wide bandwidths (90 MHz); moreover the satellite transponder essentially operates in a translational mode and hence the delay in the satellite transponder is negligible.

The major error contribution comes from the equipment delay at the ground stations. The various equipment delay contributions may be itemized as follows:

i) The time delay in the modulation circuits and the transmitter Chain at each station (δt_1, δt_2);

ii) The time delay due to the finite length of transmission line/waveguide to the transmitting antenna feedpoint (δta_1, δta_2);

iii) The time delay from the receiving antenna feed point to the parameteric amplifier input and the delay due to the transmission line/waveguide from the parametric amplifier output to the main ground station (δra_1, δra_2)

iv) The time delays in the demodulation and output circuits (δr_1, δr_2).

As mentioned earlier, any uncertainty in the time delay measurement will appear as bias in the clock offset measurement while any variability in the delays limits the accuracy.

At each earth station before the start of the experiment each day, the equipment delay was measured as part of the calibration procedure, by internal looping. A sample of the transmitter output is downconverted to the receiver frequency using a broadband mixer which contributed an insignificant amount of delay. The measured signal delay at DES was $1.28\mu s$ and for AES it was $2.8\mu s$ This AES delay included an extralength of 57m to the down converter which introduced an extra time delay of $0.57\mu s$ in the loop delay. Thus the time delay in the electronic equipment in the AES internal loop is $2.23\mu s$. The variation in the delay was less than $0.1\mu s$.

In the determination of the clock offset, the equipment time delay that enters the picture is the difference between transmitter time delay at DES plus the receiver time delay at AES and the transmitter time delay at AES and the receiver time delay at DES. Since the transmitter and receiver time delays at each station could not be measured separately, it was assumed that the total internal loop delay is equally divided between the transmitter and the receiver. Thus the net difference of the delays considered above is taken to be zero. This assumption introduces an unresolved bias of $\pm 1.0\mu s$.

The time delays introduced due to finite transmission line/ waveguide lengths at DES and AES are:

$$\delta t_1+\delta r_1 = 1.28\mu s. \qquad \delta ta_1=0.175\mu s \qquad \delta ra_1=0.2\mu s$$

$$\delta t_2+\delta r_2 = 2.8\mu s. \qquad \delta ta_2=0.364\mu s \qquad \delta ra_2=0.35\mu s$$

RESULTS

Measurements of the total time delay were made on 21 days during the two test periods in April and June 1978. Each day a maximum of 6 sets of measurements each consisting of at least 50 observations-10 centred on each minute during the 5 minute interval - have been accummulated. These data are used to determine the AES clock offset with respect to the DES clock and hence with respect to the NPL Master clock. Also because the two test periods are separated by about 6 weeks with the clock running, an opportunity is provided to determine the systematic or long - term drift rate

of the AES clock with respect to the master clock.

The clock error or offset is defined as follows: (Ramasastri et al)

Clock offset (E) = Master clock time (NPL) - User Clock Time. If the user clock lags behind the Master clock E is positive and is negative if the user clock is ahead or the master clock.

The clock offset is computed as follows. Referring to fig.2

Let τ_1 be the apparent time delay measured at AES with DES transmitting.

τ_2 be the apparent time delay measured at DES with AES transmitting.

τ_{t_1}, τ_{t_2} be the equipment delays in the DES transmit chain and receive chain respectively ($\tau_{t_1} = \delta t_1 + \delta t a_1$; $\tau_{r_1} = \delta r_1 + \delta r a_1$)

τ_{t_2}, τ_{r_2} be the equipment delay in AES transmit and receive chain respectively ($\tau_{t_2} = \delta t_2 + \delta t a_2$; $\tau_{r_2} = \delta r_2 + \delta r a_2$)

The clock offset then is given by

$$\text{Clock offset E} = \frac{1}{2}\left[(\tau_1 - \tau_2) - \{(\tau_{t_1} + \tau_{r_2}) - (\tau_{t_2} + \tau_{r_1})\}\right]$$

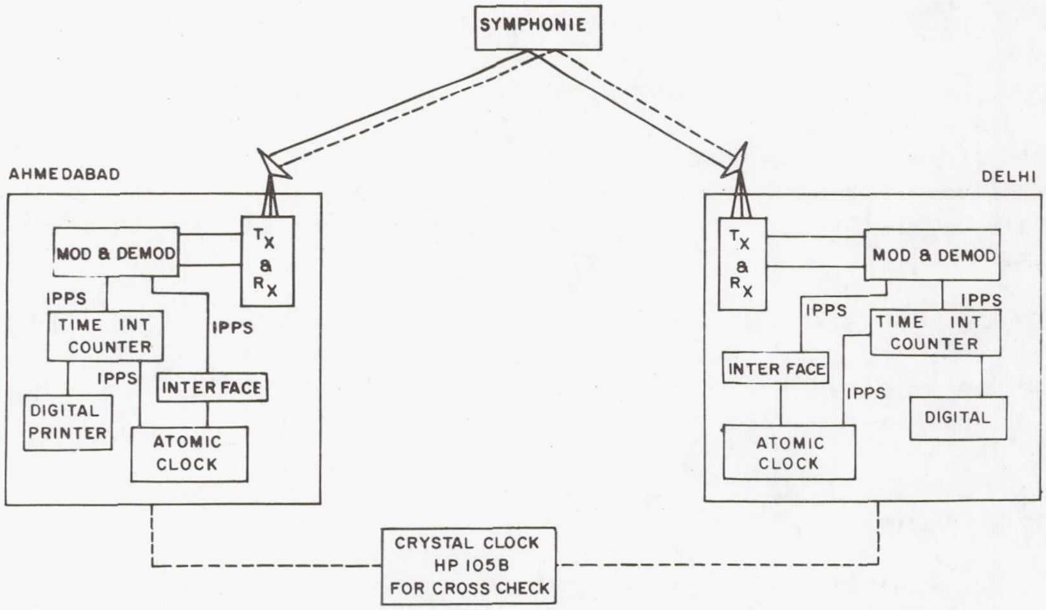
$$= \frac{1}{2}\left[\Delta\tau - \delta\tau\right] \quad \text{where } \Delta\tau = \tau_1 - \tau_2 \text{ and } \delta\tau = \left[(\tau_{t_1} + \tau_{r_2}) - (\tau_{t_2} + \tau_{r_1})\right]$$

If the equipment delay in each transmit - receive circuits is the same $\delta\tau$ is zero. However for reasons outlined in Section 4 (iii) an unresolved bias of $\pm$ 1.0μs exists.

Thus Clock offset $= \frac{\Delta\tau}{2} \pm 1.0\mu s$

As mentioned earlier, during test period each day 6 sets of observations of 5 minute duration are obtained at each station. Each sample consists of 10 observations centred on the minute. In order to take care of the systematic drift due to satellite motion, the piece-wise samples are processed on computer using a least squares fit for both the stations (Fig. 3,4). From these curves the average clock offset for each day is measured.

The corresponding standard deviation is also computed. Un-



BLOCK DIAGRAM OF TIME COMPARISION EXPERIMENT

FIG. 1

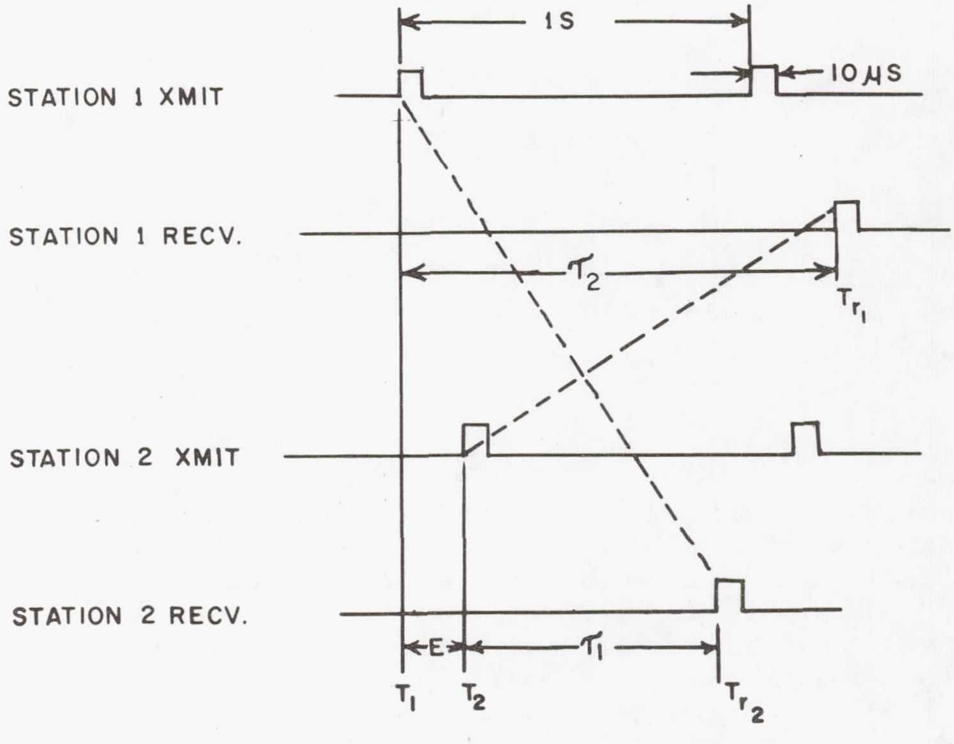


Fig. 2 TIMING DIAGRAM.

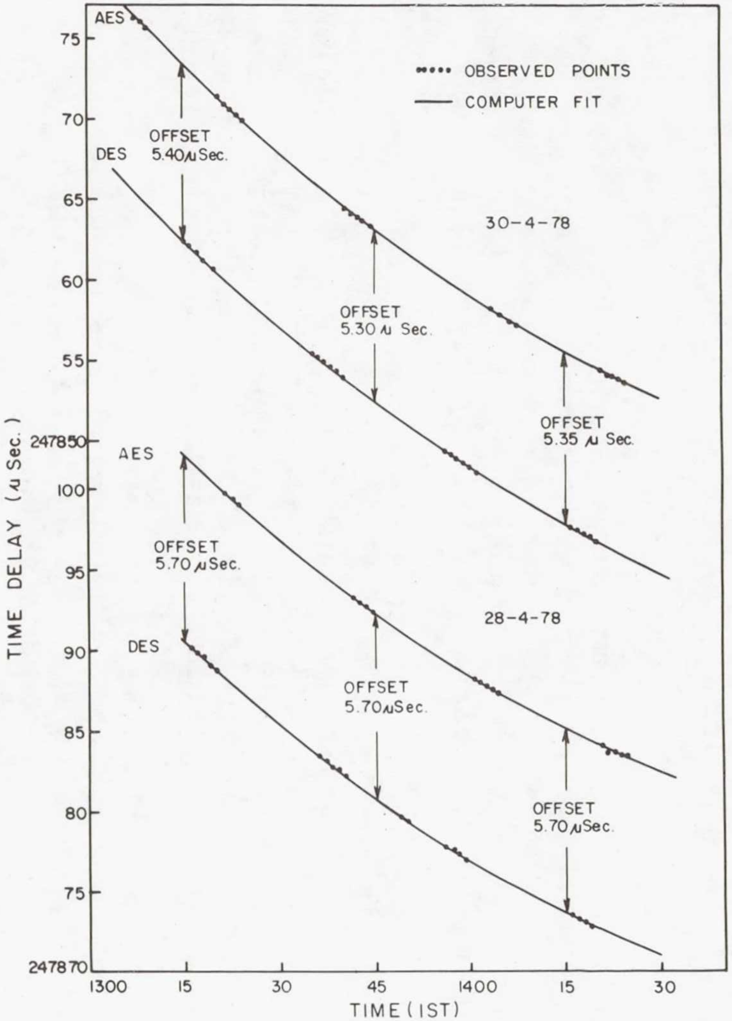


Fig. 3

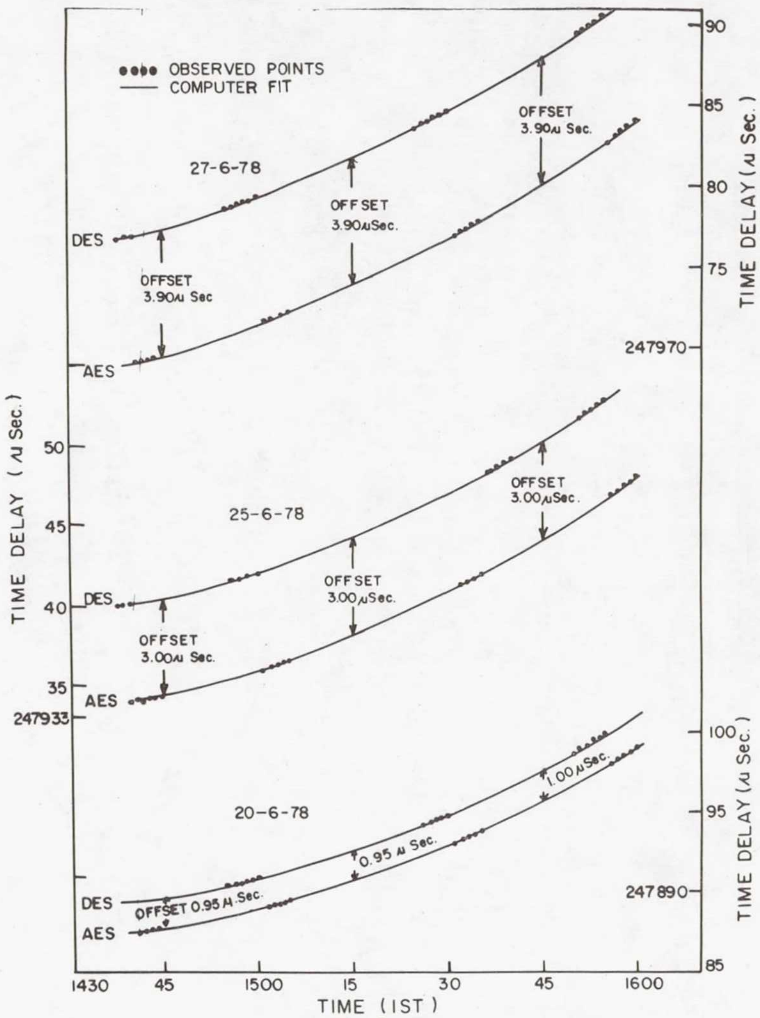


Fig. 4

Off-set between the clocks determined through computer fit in the observed time-delay points.

TABLE II a

Date	Time	Average $\Delta T/2$ (μs)	$\sigma(\Delta T/2)$	σ for fitted curves DES	AES	Counter resolution	Remarks
21.4.78	1300-1430	18.03	--	--	--	100 n Sec.	AES clock reset
22.4.78	-do-	18.10	--	--	--	100 n Sec.	--
27.4.78	-do-	7.75	--	--	--	100 n Sec.	AES clock reset
28.4.78	-do-	5.72	.034	.010	.020	100 n	--
29.4.78	-do-	5.58	.035	.016	.011	100 n Sec	-- --
30.4.78	-do-	5.35	.042	.021	.011	100 n Sec.	--

Average AES Clock Drift = .12 μSec/day

fortunately at DES the time interval counter used has only a 100 ns resolution. However a 10 ns resolution counter was used on 11 days during the tests and the standard deviation computed from these observations is taken to be representative of the capabilities of the experimental system.

The results of the measurements are summarised in table II a,b. It is concluded that the clock synchronisation is possible to a precision of .17μs using a 100 ns resolution time interval counter and to a precision of .08μs with 10ns resolution counter. The accuracy of synchronisation, taking into account all the possible delays, is about 0.5μs, with a bias of $\pm$ 1μs which is probably an overestimate.

In order to determine the drift rate of the AES clock with respect to DES clock and hence with respect to the NPL master clock, the results are plotted in Fig. 5. During the April tests the AES clock offset was + ve and it was drifting at a rate of .12μ per day up to June 16. On June 17th

TABLE II b

Date	Time	Average $\Delta T/2$ s	$\sigma(\Delta T/2)$	σ for fitted curves DES	AES	Counter resolution 100ns	Remarks
16.6.78	1430-1600	+11.21	.12	.020	.012	100n Sec.	Clock
18.6.78	-do-	.03	-	-	-	10n Sec.	Clock reset
19.6.78	-do-	-0.38	.076	.077	.014	-do-	
20.6.78	-do-	-0.95	.013	.008	.004	-do-	
21.6.78	-do-	-1.38	.288	.018	.001	Resolution 100 n Sec.	
23.6.78	-do-	-2.32	.023	.01	.002	Resolution 10 n Sec.	
24.6.78	-do-	-2.64	.097	.001	.058	-do-	
25.6.78	-do-	-3.05	.010	.005	.001	-do-	
26.6.78	-do-	-3.46	.010	.007	.008	-do-	
27.6.78	-do-	-3.92	.007	.004	.005	-do-	
28.6.78	-do-	-4.23	.076	.006	.002	-do-	
29.6.78	-do-	(-5.39)	-	-	-	-do-	Clock reset
30.6.78	-do-	(-3.49)	-	-	-	-do-	-do-

Average AES Clock Drift = .43 μSec./day

the AES clock stopped. After recommissioning the AES clock,

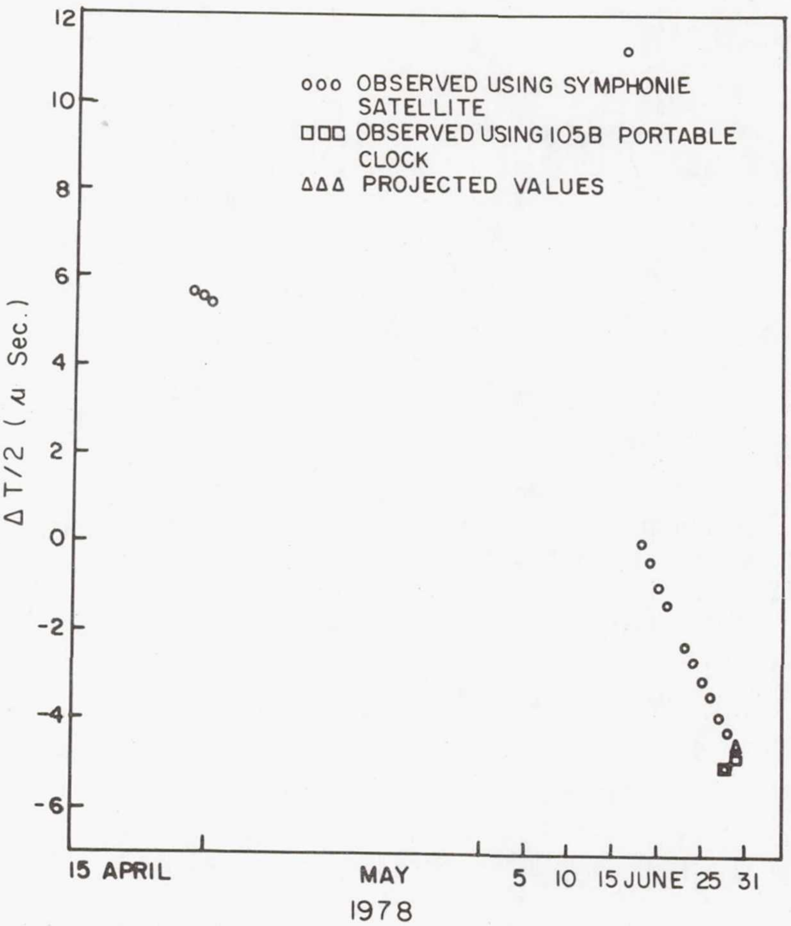


Fig. 5

Drift rate of AES clock w.r.t. DES clock

and after stabilisation it was synchronised with respect to DES clock and the offset was $\pm$ ve. During the rest of the June test period the apparent drift rate of the AES clock with respect to the DES clock was 0.5μs per day. This apparently different drift rates reconciled as follows. The DES clock has a drift rate opposite to that of AES with respect to the NPL master clock. When the AES clock offset (with respect to DES) was + ve, the apparent drift rate is smaller because the two clocks were drifting in the same direction. When the offset became - ve, the clocks were drifting in opposite directions and hence the apparent drift rate is the sum of the drift rates of the two clocks and hence is apparently much larger than during April test. The drifts of the AES and DES clocks are plotted in the same figure which establishes a uniform drift rate of the AES clock with respect to the NPL, Master Clock

Check of the Consistency of Clock Synchronisation with portable clock

The heart of the portable clock is a crystal oscillator model HP 105B. This oscillator output has been used to derive second pulses. The arrangement is also there to autosynchronise the epoch of the second pulses within a fraction of microsecond with respect to that of a cesium clock second pulses. The drift rate of the crystal clock has been studied very critically for better prediction of the clock epoch.

Before flying the clock, it is auto-synchronised with master cesium at ATA to within 0.4 microsecond and the time difference is noted in time interval counter HP 5248L. The clock is then flown to Ahmedabad by a commercial aeroplane and the AES cesium clock was directly compared with this portable clock, after about four hours time of its auto synchronisation with ATA cesium in Delhi. At Ahmedabad the offset between ATA and AES clock was found to be 1.95 μsecond with drift prediction uncertainties of $\pm$ .25μsecond (at that time according to symphonie experiment offset was 1.15μs). In the return flight the portable clock was autosynchronised with AES cesium and after roughly four hours the portable clock was again compared with ATA cesium and the offset was found to be 1.5μs (at that time the projected values of symphonie experiment was 1.32 μs).

Thus the flying clock experiment confirms the consistency of the clock synchronisation experiment via symphonie satellite to an accuracy of $\pm$ 0.25 μs.

On 28th June the difference between the flying clock experiment and the Symphonie experiment was found to be 0.8 μs which is well within the uncertainty in symphonie experiment. However, on the second day (29th June) this difference was 0.18 μs. Though both these values are within the limit of uncertainty of symphonie experiment but thedifference of 0.62 μs on consecutive two days observations may be due to some uncertainty in behaviour of crystal due to jerks while flying.

ACKNOWLEDGEMENTS

The authors wish to thank the Directors, STEP, ISRO for making the satellite Symphonie=2 available for this experiment and the staff of Delhi and Ahmedabad Earth Stations for their excellent condition and help during the course of this

experiment. The whole hearted cooperation of Telecommunication Research Centre, New Delhi through lending their sophisticated instruments whenever needed is thankfully acknowledged. The help of Indian Telephone Industries Naini who were kind enough to keep their cesium clock with NPL during the period of experiment is highly appreciated. The authors are grateful also to Indian Airlines for their kind co-operation for flying the portable clock from Delhi to Ahmedabad and back.

REFERENCES

Ramasastry,J.,B.Rosenbaum, R.D. Michelini and G.K. Kuegler: IEEE Trans. on Insts. and Meas. Vol. IM-22, p.9, 1973.

Somayajulu, Y.V. : Proceedings of the Seminar on Time and Frequency held at NPL, New Delhi during Nov. 18-20, 1976; page 509-17, 1977.

PTTI APPLICATIONS TO DEEP SPACE NAVIGATION

David W. Curkendall
Jet Propulsion Laboratory, Pasadena, California

ABSTRACT

Radio metric deep space navigation relies nearly exclusively upon coherent, i.e. 2-way, doppler and ranging for all precise applications. These data types and the navigational accuracies they can produce are reviewed. The deployment of Hydrogen maser frequency standards and the development of Very Long Baseline Interferometry (VLBI) systems within the Deep Space Network is making possible the development of non-coherent, 1-way data forms that promise much greater inherent navigational accuracy. These data, closely paralleling the observables taken with VLBI are much more sensitive to clock synchronizations, both time and frequency, and to instability during the measurement period itself than are the coherent data. The underlying structure between each data class and clock performance is charted. VLBI observations of the natural radio sources are the planned instrument for the synchronization task. This method and a navigational scheme using differential measurements between the spacecraft and nearby quasars are described.

I. INTRODUCTION AND BACKGROUND

Navigation for deep space probes has always required the acquisition of long arcs of precise doppler data to determine accurately and reliably the spacecraft orbit prior to planetary encounter. The Doppler data used is taken in the so-called coherent, or 2-way, mode. In this 2-way mode, a ground station within the Deep Space Network transmits a stable frequency reference to the spacecraft which transponds a coherent version of the received signal back to the ground station for Doppler detection (Renzetti, et al[1] and Melbourne[2]). At or near planetary encounter, the gravitational effects of the target body on the motion of the spacecraft are sufficient to produce a unique velocity profile in the Doppler that can be the dominant effect in determining where the spacecraft is. These effects usually arise too late, however, to be of benefit in the navigation process of determining and adjusting the planet-relative encounter conditions. The tracking which determines the planet-relative navigation accuracy is performed long before the

actual encounter and when the gravity forces on the probe are of little benefit in supplying the desired location information. Here the dominant information is supplied by the diurnal motion of the tracking stations as illustrated in Figure 1. A single pass of doppler tracking data will yield information about three quantities:

1) The probe's geocentric velocity by calculating the doppler shift with the diurnal effects removed.

2) The right ascension, by determining the time of the meridian crossing via observing the null in the diurnal harmonic modulation.

3) The cosine of the declination, by determining the amplitude of the harmonic modulation.

The complete spacecraft state is determined by taking additional passes of doppler data on succeeding days and combining the information from each into a full solution. When available, ranging data greatly aids this process of determining the 6-dimensional spacecraft state to the requisite accuracy. In summary:

1) range and doppler data give direct measurements of the line of sight probe distance and its time derivative,

2) the diurnal modulation of the doppler yields an indirect measurement of the right ascension, α, and declination, δ,

3) successive passes detect changes in α and δ permitting indirect measurements of $\dot{\alpha}$, $\dot{\delta}$ and provide the data volume base for noise averaging and data consistency.

Several detailed analyses of this situation have been performed - among them are: Hamilton and Melbourne[3], Curkendall and McReynolds[4] - giving the accuracy of the observable parameters as a function of doppler measurement accuracy, tracking system calibrations, probe geometry, length of the tracking pass, tracking station location errors, and random non-gravitational forces affecting the motion of the spacecraft.

The accuracy to which the indirectly measured parameters can be estimated determine the overall accuracy of the complete orbit. Shown in Figure 2 is a plot of the system accuracy at several time points from the inception of the planetary exploration era to the present time - although the list is far from complete, the margin contains a tabulation of some of the technology improvements which enabled the accuracy evolution as shown. The accuracy achievable from a given doppler and associated data calibration and processing system is seen to be a strong function of the nominal probe declination. This is because the estimate of declination is determined by observing the amplitude of the diurnal

harmonic signature, $\omega r_s \cos\delta$, and this amplitude becomes a stationary at $\delta = 0$.

The deterioration of the orbit performance at the lower range of declination can be a serious inconvenience for inner planet exploration. For example, the allowable Viking Mars arrival date space was truncated due to declinations lower than 5 deg for arrival dates past September 1976. The performance versus declination shown in the figure thus manifests itself as a mission planning constraint. Fortunately, the tempo of the geometry evolution within the inner planet system is rapid enough so that the periods of low declination and resulting loss of navigation ability are relatively short and can be usually tolerated.

The situation can be much worse for outer planet exploration. For example, the Voyager I and II Saturn encounters are both such that neither spacecraft is above 5 deg declination during the 4 months prior to the critical planetary encounter. To meet this and like situations, the dual-station planetary ranging system has been developed. Unlike the doppler system, two tracking stations are involved as is shown in Figure 3. In order to measure the troublesome declination variable these stations need to be widely separated in latitude as shown. The difference of the two range measurements is proportional to $B \sin\delta$, an observable with the right structural relationship with declination; i.e., the measurement sensitivity to declination maximizes near zero rather than becoming stationary. The declination accuracy achievable can thus approximately be given by

$$\sigma_\delta = \frac{\sigma_{\Delta\rho}}{B\cos\delta}$$

where $\sigma_{\Delta\rho}$ is the overall accuracy of the system's ability to measure the range difference, $\Delta\rho$, and B is the polar projection of the baseline. Present mechanizations require that the two range measurements be performed sequentially, rather than simultaneously, with each station obtaining a measure of the round-trip light-time between the probe and itself. The measurements can then be referred to a common epoch by either modeling the probe's motion or accumulating doppler data during the intervening period. The accuracy of this system is currently limited to a performance of approximately $\sigma_{\Delta\rho} = 4.5\sqrt{2}$m as discussed by Christensen and Siegel[5]. This measurement accuracy, working with the baseline of the tracking stations at Goldstone, CA, and Canberra, Australia, yields a declination accuracy of approximately 1 μrad over the full normal operating range of declination. It has been an iterative process, but this performance and the current Voyager mission requirements are commensurate as discussed by Jordan[6]. Figure 4 summarizes the current performance of the radio metric tracking systems for both the doppler and the ranging measurements. As shown, the doppler system provides a performance at the 0.25 μrad level for high declina-

tions. This degrades slowly until 1 μrad is reached; the dual station ranging system then provides a level 1 μrad performance for the remainder of the declination space.

II. VERY LONG BASELINE INTERFEROMETRY

Since about the beginning of the current decade, Very Long Baseline Interferometry (VLBI) systems have been developed in parallel with the coherent doppler and ranging systems just described. References 7 - 18 which span this time period, describe the development and provide analyses necessary for a detailed understanding. In this paper we shall be content to describe VLBI only in tutorial terms, emphasizing the similarities and contrasts between the VLBI data and the coherent range and doppler already discussed.

In a typical VLBI system, each of two widely spaced antennas observes a single (broad band) radio source, e.g., a quasar, simultaneously recording the received signal over a specified frequency interval. The recordings are digital in which the received voltage is digitized at the one bit level; timing information is added so that the recordings may be cross correlated later when brought to a central site. The (expected) cross correlation function can easily be shown to be approximately (Thomas[11]):

$$E\left[R(t, \Delta\tau)\right] \propto \frac{\sin \pi W\Delta\tau}{\pi W\Delta\tau} \cos \phi_1(t)$$

Where

$\Delta\tau = \tau_g - \tau_m$

$\tau_g(t)$ = geometric delay as shown in Figure 5

τ_m = a priori delay estimate inserted to bring the correlation function to near its maximum during data processing

W = bandwidth of recorded signal

$\phi_1 = \omega_1 \Delta\tau$

ω_1 = frequency at the center of the bandpass

Using typical values of 2300 MHz and 2 MHz as the rf and recorded bandwidth respectively, this correlation function goes through one complete cycle for every change in $\Delta\tau$ equal to the period of the rf frequency (<0.5 nsec). In addition it also manifests a sin x/x characteristic envelope reaching its first null at 0.5 μsec delay. These two components are called respectively "fast fringes" and the "delay function" (see Figure 6). For these same typical values, τ (or equivalently $\Delta\rho/c$,

Figure 5) can be measured directly, by adjusting τ_m so as to maximize the delay function, to on the order of 10 nsec precision (3m in light-sec). More powerful measurements can be obtained, however, in each of two conceptually different ways:

1) Observation of the source continuously over the common visibility period of the two stations (approximately 4 hours on the baselines afforded by the DSN for sources near the ecliptic), produces a continuous record of the phase of the fast fringes versus time. The record thus obtained will contain a diurnal sinusoidal modulation term due to the Earth's rotation whose phase and amplitude are parametric in the source location and baseline parameters. This is exactly analogous to the single station coherent doppler tracking except that the equatorial baseline projection and longitude play the roles of the distance off the spin axis, r_s, and station longitude, θ, respectively. The differential frequency of the two clocks replaces the geocentric velocity term observed by the coherent data.

2) Observation of the source at a second center frequency, ω_2, produces a second measurement of the fast fringe phase, ϕ_2, at a single instant of time. Then because

$$\frac{\partial\phi}{\partial\omega} = \tau_g = \frac{\phi_2 - \phi_1}{\omega_2 - \omega_1}$$

a direct measurement of τ_g can be obtained in the short time required to achieve a high S/N for the ϕ measurements (typically 10 min.). This is the "bandwidth synthesis" technique, so called because large effective bandwidths can be obtained without the need for commensurate high recording rates, and is widely used throughout the VLBI community (Rogers 9). The geometry is exactly as for the differenced range measurement already discussed (cf. Figures 3 and 4); the τ_g measurement obtained can be used directly to estimate either the baseline projection or the source location. With spanned bandwidths, $\omega_2 - \omega_1$, on the order of 40 MHz, the precision of the measurement can easily be brought to the cm level; its accuracy is dominated by other effects such as clock performance and systematic calibration errors.

When estimating source locations with method two, a second baseline is usually employed for the second component of position. An effective combination for the two baselines is to have a large polar component associated with the first and a large equatorial projection associated with the second so that they can produce largely uncorrelated estimates of α and δ.

Thus these two methods, often referred to as narrow-and wide- band VLBI respectively, have a one-to-one correspondence with the two coherent-modes, doppler and differenced ranging, normally used for spacecraft tracking. Their normal applications are duals of each other in that VLBI is usually employed to estimate the station baselines; precise source coordinates are needed to enable this. Coherent tracking is normally used to estimate the spacecraft coordinates; precise station locations are needed for this task.

Although the VLBI and the coherent tracking modes each produce observables with identical information content as just discussed, natural-source VLBI enjoys several inherent accuracy advantages over its coherent counterpart. These were discussed in some detail in a previous paper (Curkendall[19]), but briefly they include: 1) wider bandwidth, 2) more complete calibration of charged particles, 3) a ready means for calibrating electrical path delay variations in the station electronics, 4) lack of significant proper motion in the natural sources themselves, and 5) freedom from needing to model the line-of-sight motion as is required in single station doppler tracking.

There is a single major exception to the general advantages of the non-coherent data types - they suffer from a greater sensitivity to instability of the station master oscillator. This sensitivity and the comparison of it with that of the coherent data forms is treated in detail in the following section.

III. MEASUREMENT ACCURACY VS. FREQUENCY STANDARD PERFORMANCE

The two measurement classes just discussed, coherent measurements and non-coherent VLBI measurements differ markedly by the manner in which the station's master frequency standard departures from ideal enter and corrupt the measurements.

In this section the four data types:

Data Class / Data Type	Two-Way Coherent Measurements	One-Way Non-Coherent Measurements
Narrowband	Doppler (1)	Narrowband VLBI (2)
Wideband	Differenced Ranging (3)	Wideband VLBI (4)

will each be analyzed and their sensitivity to clock performance charted.

Case 1 - Coherent Narrowband (Doppler) Data

Consider the (highly) schematic diagram of a typical coherent doppler and ranging system implementation as shown in Figure 7. Counted, or integrated, doppler is obtained by broadcasting a stable reference to the spacecraft which coherently transponds the received carrier back to the same station for comparison with the original transmitted frequency; the difference or doppler frequency is integrated by means of a counter as shown. Assume for the purposes of illustration, that a unit step in frequency error occurs for a short period of time as shown in Figure 8. This will enter the doppler extractor and be integrated to yield immediately a buildup of range error, $\Delta\rho_\varepsilon$. It will also be transmitted to the spacecraft and return a round-trip light time, τ, later and re-enter the doppler extractor, this time in the opposite sense - $\Delta\rho_\varepsilon$ will return to zero. If the doppler system has been tracking the spacecraft for T seconds, the accumulated effect of the time history of $\Delta f(t)$ is readily seen to be

$$\Delta\rho_\varepsilon \begin{array}{l} = \dfrac{c}{2f}\displaystyle\int_0^T \Delta f(t) - \Delta f(t-\tau)\, dt = \\[2ex] = \dfrac{c}{2f}\displaystyle\int_{T-\tau}^T \Delta f(t)dt - \int_{-\tau}^0 \Delta f(t)dt \end{array} \qquad (T \gtrsim \tau) \qquad (1)$$

where c is the speed of light.

That is, in a data point measured at T, frequency standard performance during the first and last τ seconds counts, everything else cancels out. The factor of 2 appears so that $\Delta\rho_\varepsilon$ is the error in the one-way range as measured by the two-way instrument. It is useful to design expressions which permit calculating $\Delta\rho_\varepsilon$ assuming Δf is

i) a white noise process

ii) linear with time

i) White Frequency Noise

With the white noise assumption, the two integrals in (1) are clearly independent, the variance of $\Delta\rho_\varepsilon$ is then

$$\sigma_{\Delta\rho_\varepsilon} = \frac{c}{2}\sqrt{2\int_0^\tau\int_0^\tau h_o\ \delta(u-v)\,du\,dv} = \frac{c}{2}\sqrt{2\ h_o\ \tau} \tag{2}$$

where h_o is the spectral density of the white frequency $\Delta f/f$ process.

ii) Linear Frequency Drift

$$\text{Assume}\quad \frac{\Delta f(t)}{f} = kt$$

$$\text{Then from (1),}\quad \sigma_{\Delta\rho_\varepsilon} = \frac{c}{2}\,k\ T\ \tau \tag{3}$$

where k is re-interpreted to be the standard deviation of the drift. Thus for frequency variations rapid relative to τ, the error builds with $\sqrt{\tau}$ and is independent of the tracking time; for frequency variations slow relative to τ, the error builds as the product, $T\tau$.

Case 2 - Non-Coherent, Narrowband (VLBI) Data

The non-coherent case is even more straightforward. Here the error in the measurement is easily seen to be proportional to the difference in the oscillators' frequencies at the two Earth-based stations integrated over the observational interval:

$$\Delta\rho_\varepsilon = \frac{c}{f}\int_0^T f_1 - f_2\ dt \tag{4}$$

where f_i = instantaneous frequency at the ith station.

The Δf notation used in (1) is dropped here to emphasize that the measurement is sensitive to more than just the change in frequency over τ or even T, it is sensitive to the "knowability" of the frequency difference. The initial frequency offset is large enough so that in any VLBI experiment it must be considered an unknown and solved for - indeed, the determination of $f_1(0) - f_2(0)$ is often the reason for the

VLBI experiment. It becomes convenient, then, to re-define the problem and focus on a modified error term as

$$\Delta\rho'_{\varepsilon} = \frac{c}{f}\int_{0}^{T} \Delta f_1 - \Delta f_2 \, dt \tag{5}$$

where Δf_i is understood to be the departure in the frequency from $f_i(0)$. Then

$$\sigma_{\Delta\rho'_{\varepsilon}} = \sqrt{2}\; c\left(\frac{\Delta f}{f}\right)_T T \tag{6}$$

where $\left(\frac{\Delta f}{f}\right)_T$ is the familiar two-sample

Allan Variance over smoothing time, T. Equation (4) is already in what is usually the most convenient form. For comparative purposes, however, it is interesting to recompute the effects arising from the white noise and linear drift models used earlier.

i) White frequency noise

$$\sigma_{\Delta\rho_{\varepsilon}} = \sqrt{2c^2 \iint_{0}^{T} h_o \; \delta(u-v)\,du\,dv} = c\sqrt{2h_o T} \quad \text{(c.f.2)} \tag{7}$$

ii) Linear Frequency Drift

$$\frac{\Delta f_1 - \Delta f_2}{2} = \sqrt{2}\,kt \tag{8}$$

$$\sigma_{\Delta\rho'_{\varepsilon}} = \frac{\sqrt{2}\,k\,T^2}{2} \quad \text{(c.f.3)}$$

Case 3 - Differenced Coherent (Ranging Data)

A coherent range measurement is essentially a measurement of the round-trip light time itself. The clock error introduced in such a measurement is thus the absolute frequency error integrated over the light time. A differenced ranging measurement is sensitive to the integrated frequency differences, i.e.,

$$\Delta\rho_{\varepsilon} = \frac{c}{2f}\int_{0}^{\tau} f_1 - f_2 \, dt \tag{9}$$

For precision differenced ranging measurements, this frequency difference (or more precisely, the error in the knowledge of the difference)

must be held to within strict limits. For example, at the distance of Saturn, $\tau \approx 10^4$ sec., $|f_1 - f_2|$ must be $< 3 \times 10^{-13}$ for a $\Delta\rho_\varepsilon$ of less than .5 m. The implied frequency synchronization must be accomplished with traveling clocks or, as in more frequently the case, with a VLBI observation session whose object is the determinations of $f_1 - f_2$.

The error, then, in applying the synchronization to a differenced range measurement can be roughly predicted as

$$\sigma_{\Delta\rho_\varepsilon} = c \left(\frac{\Delta f}{f}\right)_{\Delta t} \tau \tag{10}$$

where Δt is the time between synchronization and application. This expression assumes that the synchronization operation has an associated error much less than $(\Delta f/f)_{\Delta t}$. This is not always the case and the VLBI method for synchronization brings up an interesting interplay between clock performance and its measurement. In an ideal experiment where all parameters influencing the interferometer phase (geometry, media transmission effects, etc.) are known save the clock offset, $f_1 - f_2$, itself, the clock performance during the experiment of duration, T, will contribute an error

$$\sigma_{f_1 - f_2} = \frac{\sigma_{\Delta\rho_\varepsilon}}{c\ T} = \sqrt{2}\left(\frac{\Delta f}{f}\right)_T \tag{11}$$

where $\sigma_{\Delta\rho_\varepsilon}$ is given by (6).

Case 4 - Non-Coherent Wideband (VLBI) Data

Here the measurement error is proportional to the clock time offset at the time of the measurement. Once again, this parameter must be measured periodically in order that the knowledge of it can be held to within reasonable limits. The expression for the error using the same nomenclature as the first three cases would be

$$\Delta\rho_\varepsilon = \frac{c}{f}\int_{"-\infty"}^{T} f_1 - f_2 \, dt \tag{12}$$

where "$-\infty$" is understood to be the time of the last clock epoch synchronization operation. As before, if the synchronization operation is assumed accurate, the predicted standard deviation of (12) is readily seen to be

$$\sigma_{\Delta\rho_\varepsilon} = c\left(\frac{\Delta f}{f}\right)_{\Delta t} \Delta t \tag{13}$$

where $\Delta t = T -$ "$-\infty$".

Evaluation of (13) for even Hydrogen maser stabilities ($\Delta f/f \approx 10^{-14}$) discloses that Δt cannot exceed 1/3 day if decimeter level measurements are sought. Because of this high sensitivity, it is common practice to include provision for both clock epoch and clock frequency synchronization integral to any natural source VLBI experiment. Equation (13) has relevance strictly only when envisioning a series of natural source measurements for clock synchronization whose results are applied to subsequent (or earlier) spacecraft tracking.

Table I summarizes the four cases just discussed and repeats the four basic sensitivity equations. Table II is an attempt to tabulate the expected metric errors arising in the same four cases in terms of the two-sample Allan Variance. There are several approximations used in writing the expressions shown and this table should be viewed more as an ordered collection of the principles discussed here rather than a set of rigorous relationships.

The non-coherent VLBI measurements are thus much more demanding of clock performance than their coherent counterparts. Indeed, the relative immunity of coherent doppler and range to clock variations enabled reasonable performance at small light-time distances ($\leq$ lunar distance) even with the crystal oscillators that were employed in the early 60's. With current rubidium standards ($\Delta f/f < 10^{-12}$/day) good performance can be achieved throughout the terrestrial planet space. In contrast, VLBI errors are dominated by clock effects when rubidium and even cesium ($\Delta f/f < 10^{-13}$/day) standards are used. The introduction of the Hydrogen maser with drifts better than 10^{-14}/day has in large part prompted the current interest in VLBI systems. At this clock performance level, the VLBI accuracy estimates given in the next section can be achieved and the non-coherent measurement class can successfully compete with its coherent counterpart.

IV. VLBI AS A NAVIGATION TOOL

The maturing VLBI technology can be applied to the spacecraft navigation problem in each of two conceptual ways: 1) calibration of the DSN for use as an otherwise conventional radiometric network, and 2) direct spacecraft signal tracking with the VLBI data acquisition and processing systems.

VLBI for DSN Calibrations

In July of '79, the first operational VLBI system (as contrasted with the R&D systems which have produced the results discussed above) will begin taking routine measurements on the California-Spain and California-Australia baselines. Initially this system, operating in conjunction with a developed precision source catalog, will have the capability of operating with an overall accuracy at the 0.05 µrad level. This capa-

bility will be exploited to calibrate the UT and PM variations, refine the knowledge of baseline vectors, and determine the relative station clock frequency offsets. The UT/PM calibration improves the knowledge of the effective station location coordinates for both range and doppler; the baseline solutions directly improve the differenced range data. Both of these error sources will be controlled to the 50 cm level as compared to their current 1-2 m effective levels. The interstation clock frequency calibration will be accurate to better than 1×10^{-13}, bringing the clock's contribution to the differenced range error (Eq. (10)) to well under 50 cm even at 10 AU.

Direct VLBI Navigation

The spacecraft can be treated as another VLBI radio source, albeit one with proper motion. Present spacecraft do not emit signals with bandwidths wide enough to really be considered wideband sources, however. Instead, the power is centered within a few MHz of the carrier, effectively enabling tracking at only the carrier center frequency so that the fast fringes may be observed but bandwidth synthesis at beneficial accuracies cannot be performed. Thus, only narrowband VLBI can be made available for already launched spacecraft. Save the one feature of the wide bandwidth, all of the inherent advantages of the non-coherent data class can be capitalized on, however.

The favored operational procedure in narrowband spacecraft VLBI is to track the spacecraft in close conjunction with the tracking of a nearby quasar ($\sim <5$ deg away in angular measure) whose source location has already carefully been determined. In this procedure, known as ΔVLBI, the two antennas track the spacecraft for a few minutes and then quickly align on the quasar. This is repeated for the entire duration of the 4 hour tracking pass and the fringe phase history differential between the spacecraft and the source is constructed. The differences in α and δ between the two sources are then estimated. This procedure has the advantage that most of the systematic error sources, being nearly common to both the spacecraft and quasar phase histories - baseline coordinates, instrumentation calibration, neutral and charged particle media effects, and clock imperfections - are diminished by the differencing operation. The precise natural source location estimate is quickly transferred to the spacecraft by this technique.

It is relatively simple to alter the broadcast spacecraft waveform to make it suitable for wideband VLBI tracking. Modulating a high frequency subcarrier with a signal generated from a noise diode could produce noise channels suitable for bandwidth synthesis with characteristics nearly identical to that being received from the quasar. A far better technique is to transmit a specific ranging code generated onboard the spacecraft and correlate the received spacecraft signal at each station against a locally generated model of that same code in a manner similar to that used in a conventional 2-way ranging machine.

In contrast to the current ranging technique, the signal would not be demodulated and tracked via the closed loop receiver, however. Rather, the spacecraft signal would follow the same rf chain as the quasar signal, phase calibration signals would be introduced, and detection would take place after digitization so as to preserve the near perfect commonality with the natural radio source tracking essential for careful system calibration. Each station executes this procedure, a one-way range is calculated, and the difference taken. This difference contains the differential station clock epoch error which must be carefully calibrated with conventional VLBI. This calibration, along with UT/PM calibration, can be performed via either an earlier multiple-source VLBI run and applied to the Differential One-Way Ranging (DOR) data or the DOR data can be obtained in conjunction with the track of a single very nearby quasar in a manner similar to the narrowband ΔVLBI method already described. Which method is employed operationally would be determined by the availability and strength of a nearby quasar, the number of spacecraft that need to be tracked, the availability of transmission lines for the VLBI data, the time criticality of the spacecraft results, and the ability of the clocks to hold an epoch synchronization.

An extensive error analysis and discussion of the wide-band VLBI, or DOR, technique was presented by Melbourne and Curkendall in an earlier paper[20]. As a result of that and subsequent analysis, it is felt that a confident prediction of the performance of such a system can be made at the 0.05 μrad level. The narrowband ΔVLBI can perform at similar levels for spacecraft at high declination, but would degrade with the $1/\tan\delta$ characteristic inherent for all narrowband tracking.

These accuracies, along with the accuracy of the coherent tracking earlier discussed, is summarized in Figure 4 . In contrasting the wide and narrowband approaches, the constant performance of the wide-band method with declination is of the utmost importance. In addition, the tracking time required for a complete observation is much smaller for the wide-band system since no diurnal signature need be observed. Observations from two baselines are required however. The narrowband approach suffers from a greater sensitivity to systematic error sources so only the ΔVLBI mode should be used to meet a precision requirement. The wideband is more immune and offers greater tracking strategy flexibility. It is thus less dependent on the existence of a nearby quasar. This could be important for some applications since at the present time, few suitable sources have been found in the portion of the celestial sphere where the galaxy partially obscures extragalactic observations (Preston et al[21]). This results in two bands along the ecliptic plane, each measuring about 30 deg in extent, nearly devoid of suitable sources. Finally, the post correlation data processing is more difficult for the narrowband ΔVLBI data. The key to the accuracy is the accurate construction of the accumulated phase delay change over the 4 hour period for both the spacecraft and the quasar. Care must be taken to ensure that the phase is extrapolated properly to within one

rf cycle (13 cm at S; 3.5 cm at X-band) during the time periods when the antennas are moving or are on the opposite source.

On a more positive note, narrowband ΔVLBI can be very useful for tracking existing spacecraft and offers a greater sensitivity for the detection of proper motion over short time periods (< 4 h).

Sensitivity of Differential VLBI to Clock Performance

Differential-, or Δ-VLBI, is intermediate in clock sensitivity between the coherent or non-coherent data forms discussed in Section III and for that reason deserves special treatment in this concluding sub-section.

This sensitivity is generally proportional to the angular separation between the natural source and the spacecraft. For example, in narrowband ΔVLBI, when the sources are both within a single beamwidth of the antennas, the normal switching between the two as described above is not necessary, both sources are tracked all the time and the effect of clock variations cancels out. All that is needed in this case are clocks with coherence times long enough to detect the signals (∿ a few minutes). Beyond a single beamwidth separation, antenna switching is necessary and the individual spacecraft and quasar records will have complimentary data outages as shown in Figure 9. For separations just greater than the single beamwidth limitation the data differencing operation would difference the two data streams as close together in time as possible or one full cycle time, denoted as Δt in the diagram. Consider for the moment, that the clock instability is the sole error source, the measured Δρ for the quasar would be:

$$\Delta\rho_{mq}(T) = \Delta\rho_q(T) + \frac{c}{f}\int_0^T f_1 - f_2 \, dt \tag{14}$$

An identical expression for the spacecraft measurement, $\Delta\rho_{m\ s/c}(T)$ can also be written. Their difference with the time shift, Δt is

$$\begin{aligned} \Delta\Delta\rho_m(T, \Delta t) &= \Delta\rho_{mq}(T + \Delta t) - \Delta\rho_{m\ s/c}(T) \\ &= \Delta\rho_q - \Delta\rho_{s/c} + \frac{c}{f}\int_T^{t+\Delta t} f_1 - f_2 \, dt \end{aligned} \tag{15}$$

It is useful to calculate the expected rms value of the clock induced error for the white frequency noise and linear drift examples used earlier.

i) White Frequency Noise

$$\sigma_{\Delta\Delta\rho_\varepsilon} = c\sqrt{2h_o \ \Delta t} \qquad \text{(c.f.2\&7)} \tag{16}$$

ii) Linear Frequency Drift

$$\sigma_{\Delta\Delta\rho_\varepsilon} = \frac{ck}{2}\left(\Delta t^2 + 2T\Delta t\right) \qquad \text{(c.f.3\&6)} \tag{17}$$

In comparing these expressions with those derived earlier, note that if the switching time is shorter than the spacecraft round trip light time (Δt can be on the order of 10 minutes, τ is often several hours), the ΔVLBI data is actually less sensitive to the white frequency noise than is two-way counted doppler. The expression for the linear drift model is a good illustration that the sensitivity to more systematic clock errors is intermediate to that of the coherent and non-coherent forms discussed in Section III.

For larger source separations, the situation is somewhat more complex. As the angle increases, it soon becomes apparent that if both sources are tracked over the same time periods (e.g. $t = 0$ to T) except for the alternating outages, the cancellation of other error sources will not be nearly so complete as if better strategies had been employed. For example, suppose that the sources have equal declinations but differ in their right ascensions and that the dominant error source is the tropospheric scale height. It should be clear that in this circumstance, the best strategy would be to track each source over the same range of hour angles and stagger the time intervals as shown in Figure 10.

In this circumstance, the effective Δt would grow to a much larger value than is strictly needed for the switching time operation as is illustrated in the figure. More generally, the global tracking scenario and the effective Δt must be chosen to minimize the sum of all the error sources, a problem beyond the scope of this short discussion. Once chosen however, the clock's contribution to the accumulated error can be written directly in terms of the Allan Variance as

$$\sigma_{\Delta\Delta\rho_\varepsilon} = \sqrt{2}\, c \left(\frac{\Delta f}{f}\right)_{\Delta t} \Delta t \tag{18}$$

In wideband ΔVLBI, the tracking strategy is simplified to a single observation of each source. The clock's contribution to the error in the difference of the two resulting time delay measurements are 1) the error suffered internal to the measurement itself, and 2) the clock offset drift in between the two measurements, spaced Δt apart. The determination of the first of these is beyond the scope of this discussion (but should be small), the second is given by (18) just as in the narrowband case. The real difference between the narrow and wideband cases is that the narrowband transformation to right ascension and declination estimates is itself more sensitive to a time delay error change (by about a factor of 5 at high declination, growing gradually worse as declination is reduced) than is the wideband transformation sensitivity to time delay error.

A single numerical example should serve to put these relationships in perspective. During the 1979 Voyager encounters with Jupiter, a narrowband ΔVLBI demonstration is planned whose accuracy goal is set at the .05 μrad level. A single quasar, OJ287, is being used as the reference natural source throughout the several month experimentation period; the appropriate Δt is as large as 45 min. The error control needed to achieve this accuracy is approximately 7 cm of range change error during the 4 hour integration period. If 1 cm is the clock's allocation, the two-sample Allan Variance clock performance required is (from (18):

If $\quad \sigma_{\Delta\Delta\rho_\varepsilon} = .01\text{m}$

$$\Delta t = 45 \text{ min.}$$

$$\frac{\Delta f}{f} = \frac{\sigma_{\Delta\Delta\rho_\varepsilon}}{\sqrt{2}\, c\, \Delta t} = 8.8 \times 10^{-15}$$

If wideband ΔVLBI were possible, the same accuracy level could be achieved with a total error budget of about 40cm. If 1/7 of this were allocated to the clock, the $\Delta f/f$ specification could be relaxed to 5×10^{-14}. Alternately, and probably of more practical importance, a

clock operating at 8.8 x 10^{-15} could be used to lengthen the time between epoch calibration and spacecraft use; i.e., the 45 min. could be stretched to several hours.

TABLE I: Comparison of Data Type Sensitivity to Station Master Clock Errors

Data Class / Data Type	Two-Way Coherent Measurements	One-Way Non Coherent Measurements
Narrowband	<u>Doppler</u> Clock Frequency Drift During Pass Integrated Over Round Trip Light Time to S/C $\Delta\rho_{\varepsilon} = \frac{c}{2f}\left[\int_{T}^{T+\tau} f(t)\ dt - \int_{o}^{\tau} f(t)dt\right]$	<u>Narrowband VLBI</u> Interstation Frequency Difference Integrated Over Tracking Pass Duration $\Delta\rho_{\varepsilon} = \frac{c}{f}\int_{o}^{T} f_1 - f_2\ dt$
Wideband	<u>Differenced Ranging</u> Interstation Clock Frequency Difference Integration Over Light Time $\Delta\rho_{\varepsilon} = \frac{c}{2f}\left[\int_{o}^{\tau} f_1 - f_2\ dt\right]$	<u>Wideband VLBI</u> Interstation Clock Epoch Error at Measurement Time $\Delta\rho_{\varepsilon} = \frac{c}{f}\int_{``-\infty"}^{T} f_1 - f_2\ dt$

f_i = instantaneous frequency of i^{th} station clock

T = Tracking Pass Duration

τ = Round Trip Light Time to S/C

NOTE: All errors are given in range difference or integrated range-rate (m).

Table II.

Approximate Evaluation of Expected Metric Error, $\sigma_{\Delta\rho_\varepsilon}$, In Terms of the Two Sample Allan Variance $\left(\frac{\Delta f}{f}\right)_t^2$.

Data Type \ Data Class	$\sigma_{\Delta\rho_\varepsilon}$ Two-Way Coherent Measurements	$\sigma_{\Delta\rho_\varepsilon}$ One-Way Non-Coherent Measurements
Narrowband	$\frac{\sqrt{2c}}{2}\left(\frac{\Delta f}{f}\right)_\tau^{*}\tau$ *only strictly true for white phase and frequency noise processes	$c\left[\frac{\Delta f_{\varepsilon_s}}{f} \oplus \sqrt{2}\left(\frac{\Delta f}{f}\right)_{\Delta t}^{+} \oplus \sqrt{2}\left(\frac{\Delta f}{f}\right)_T\right]T$
Wideband	$\frac{c}{2}\left[\frac{\Delta f_{\varepsilon_s}}{f} \oplus \left(\frac{\Delta f}{f}\right)_{\Delta t}^{\dagger}\right]\tau$ †an approximate interpretation of Allan Variance	$c\left[\Delta\tau_{\varepsilon_s} \oplus \frac{\Delta f_{\varepsilon_s}}{f}\Delta t \oplus \left(\frac{\Delta f}{f}\right)_{\Delta t}\Delta t\right]$

where

c = speed of light

$\left(\frac{\Delta f}{f}\right)_t$ = two sample Allen Variance evaluated at t smoothing time

τ = round trip light-time to coherent transponder

T = time from beginning of $\Delta\rho$ integration

Δf_{ε_s} = error from frequency synchronization operation

$\Delta\tau_{\varepsilon_s}$ = error from clock epoch synchronization operation

Δt = time from clock synchronization to T

$\oplus$, an operator impling addition of errors in the rms sense

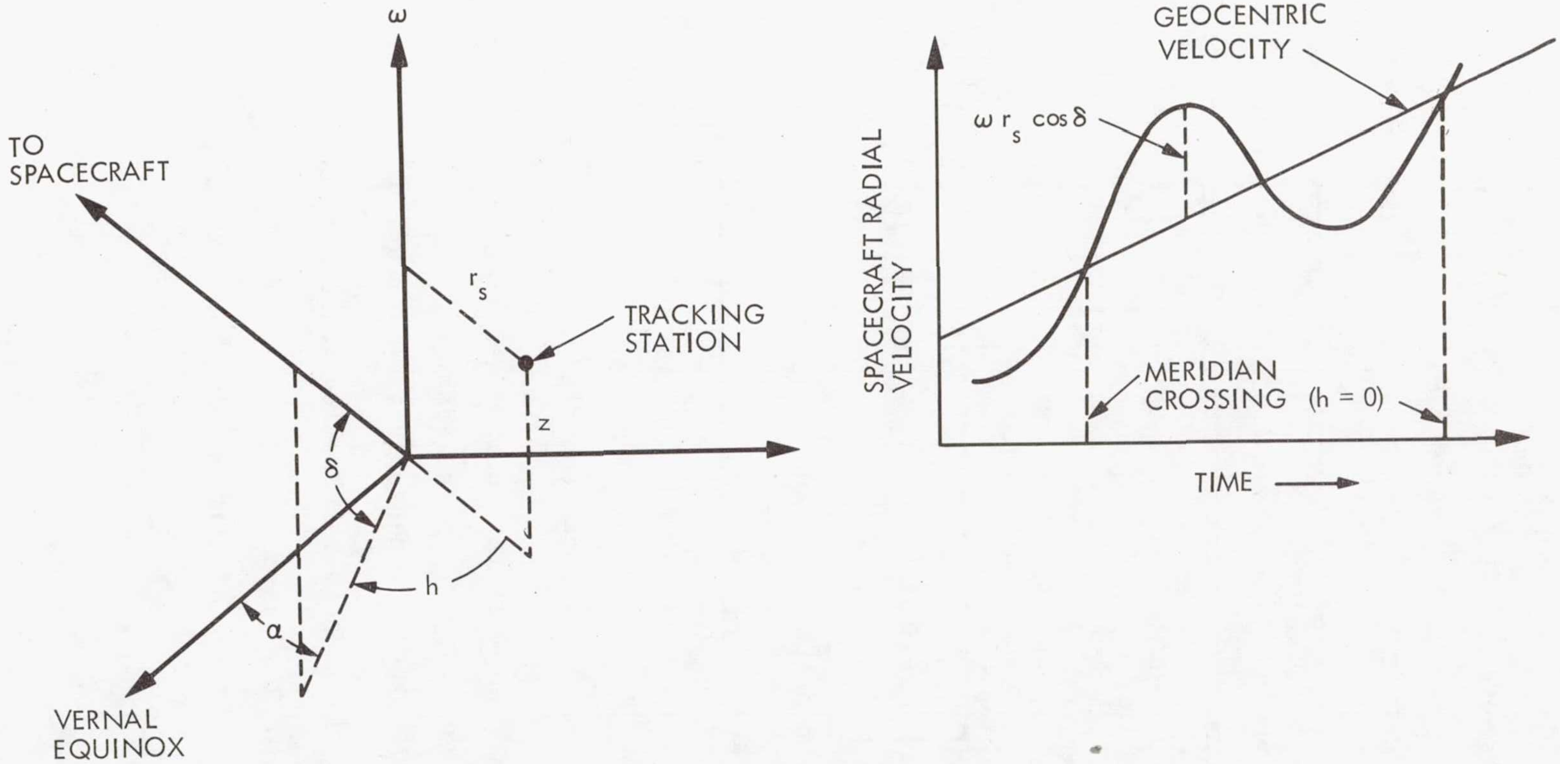


Figure 1. The Primary Information Source for Pre-Encounter Navigation Arises from the Diurnal Motion of the Tracking Station.

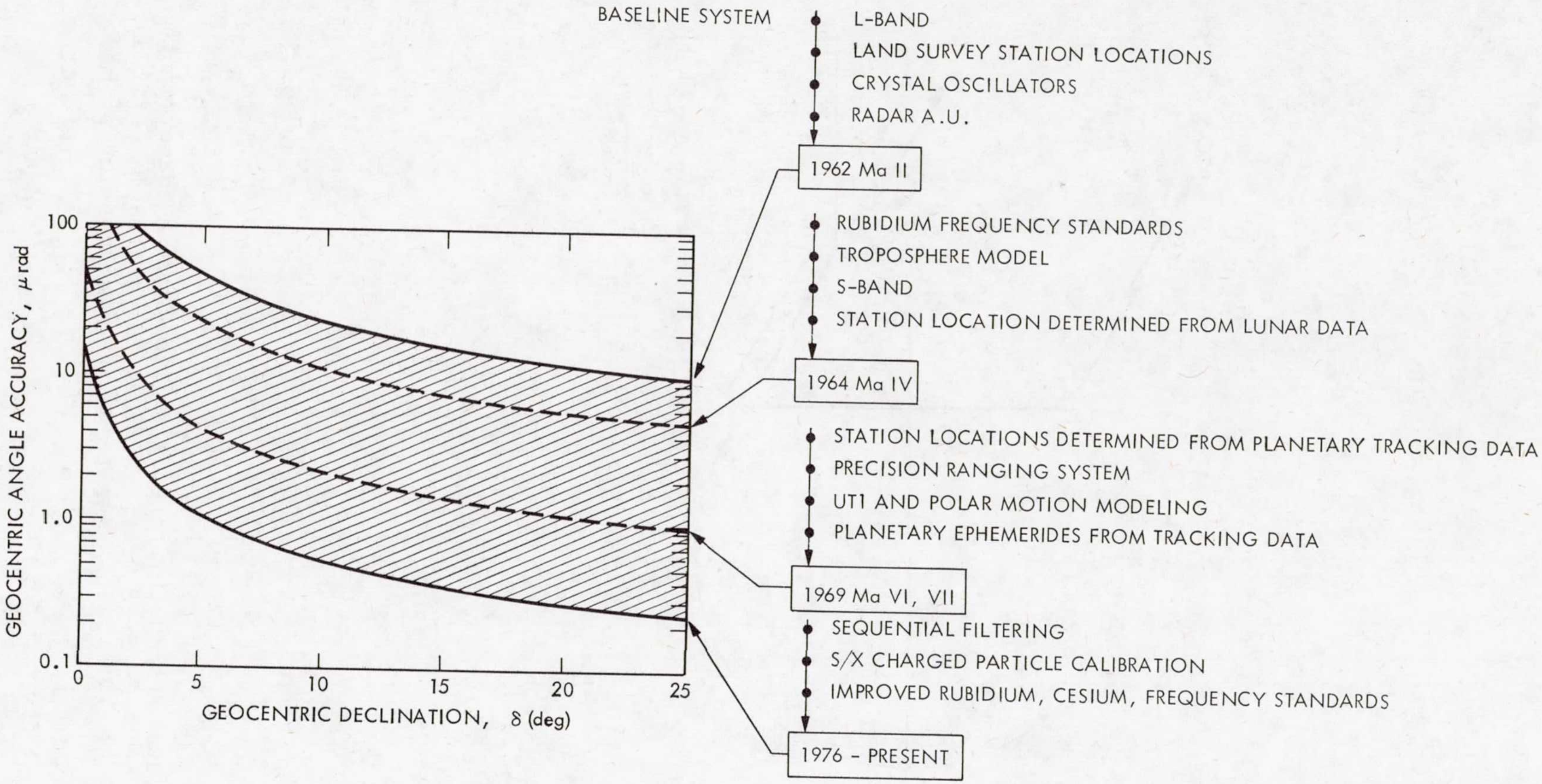


Figure 2. Doppler System Performance versus Time and Spacecraft Declination

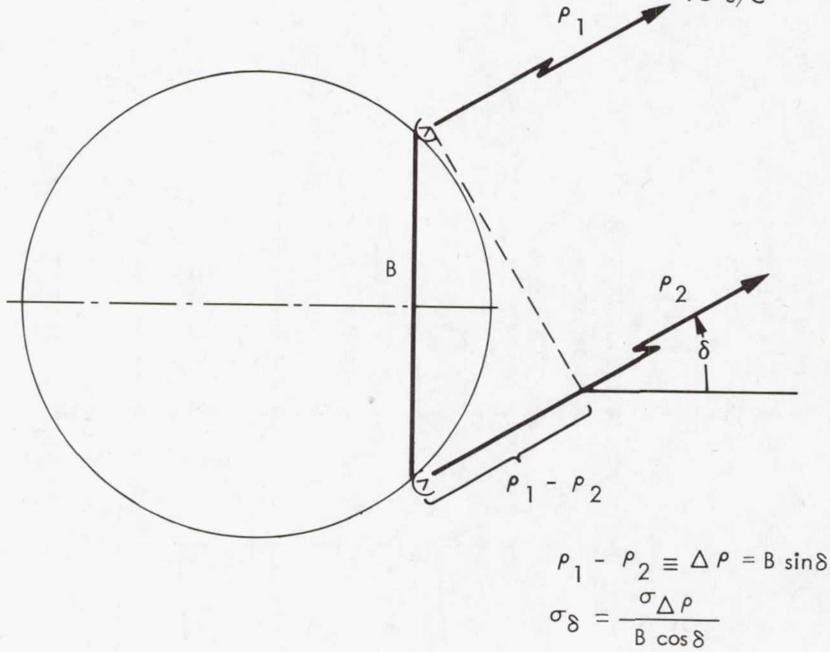


Figure 3. The Difference of Two Ranging Measurements from Stations Widely Separated in Latitude can Accurately Measure Declination of Probe

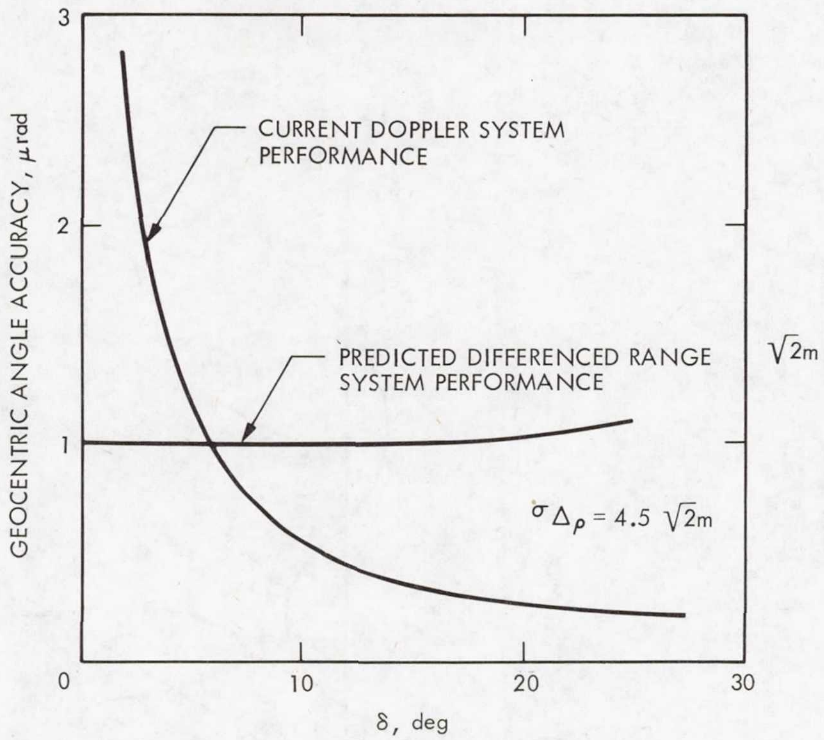


Figure 4. Accuracy Performance of Doppler and Ranging System versus Nominal Probe Declination.

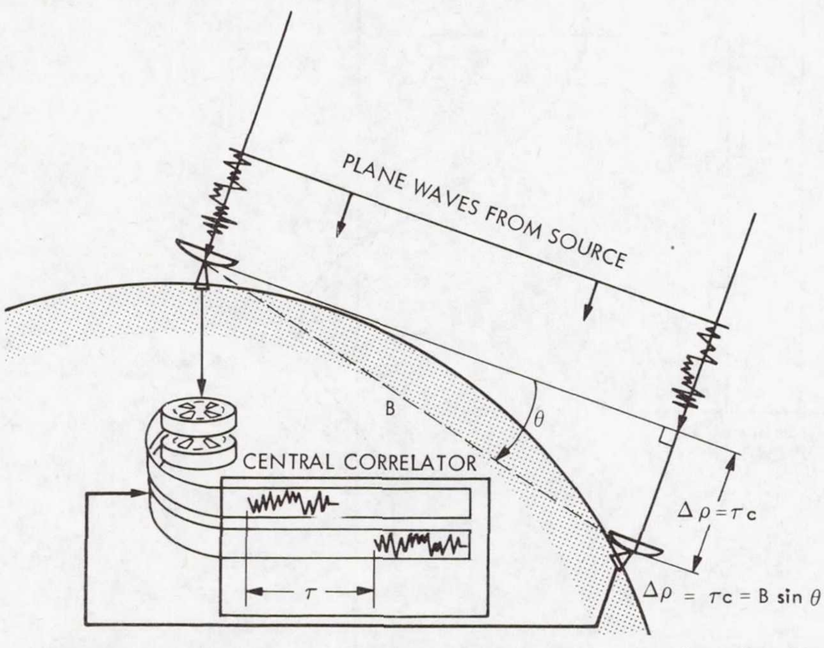


Figure 5. Wideband VLBI can Estimate Source Location by Measuring the Differential Time of Arrival.

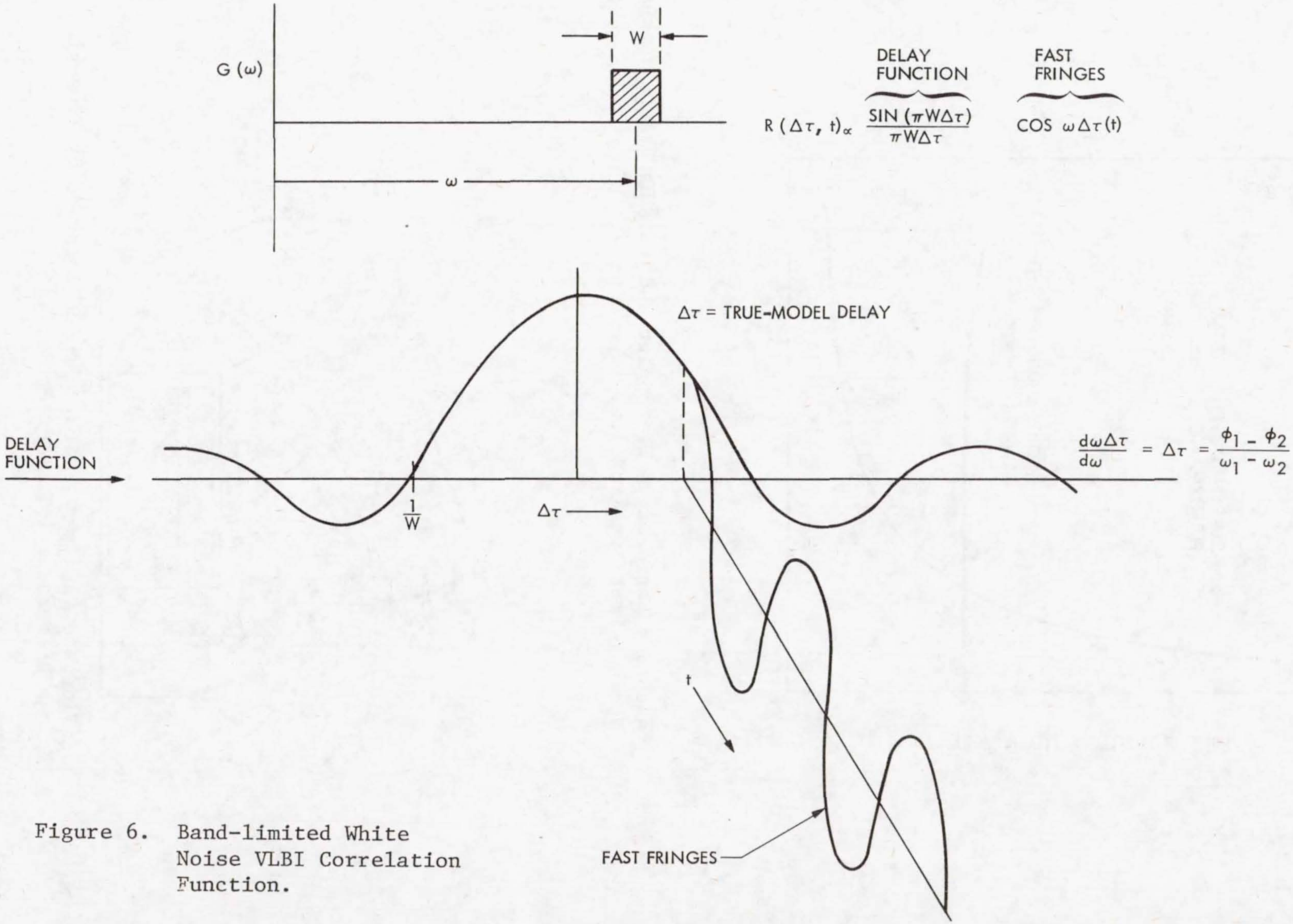


Figure 6. Band-limited White Noise VLBI Correlation Function.

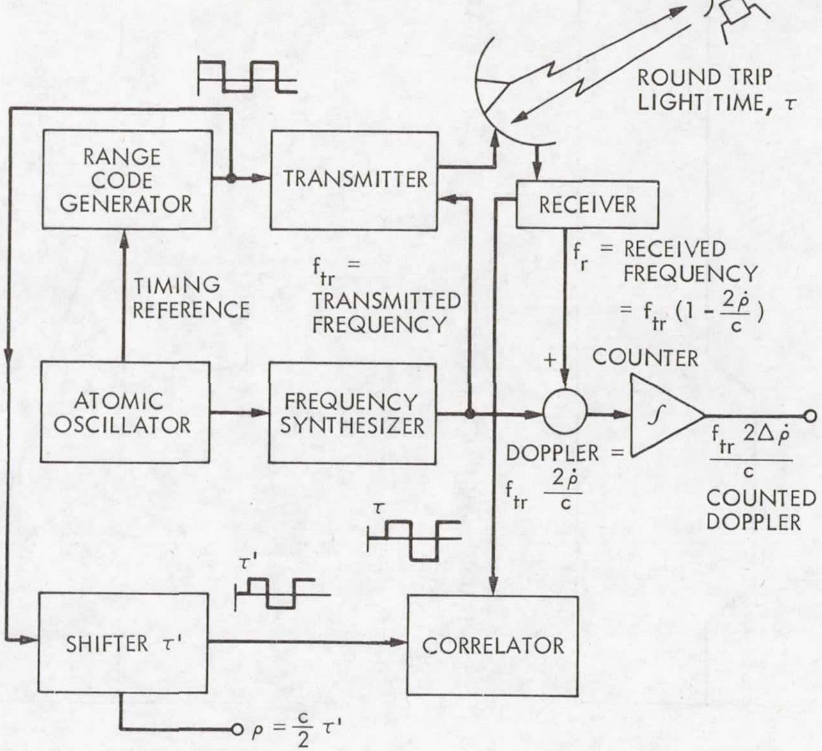


Figure 7. Coherent Doppler, Range Ground Tracking System

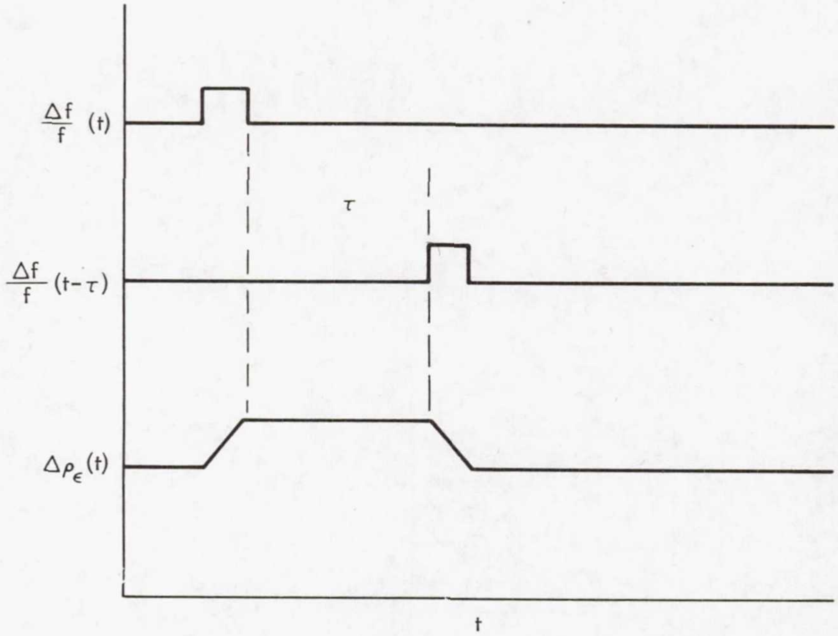


Figure 8. Counted Doppler Response, $\Delta\rho_{(-)}$ To a Unit Step of $\Delta f/f$.

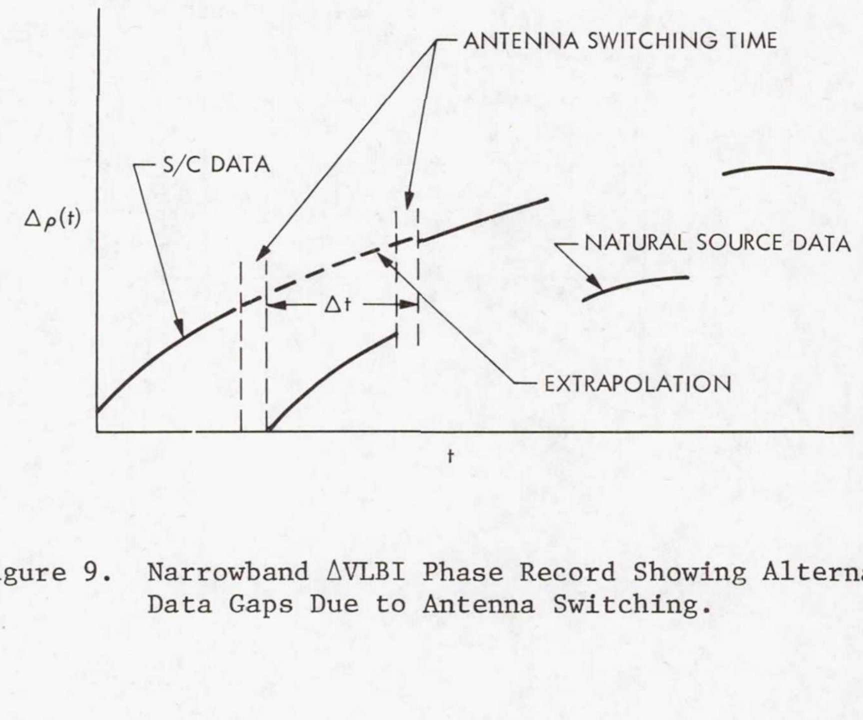


Figure 9. Narrowband ΔVLBI Phase Record Showing Alternate Data Gaps Due to Antenna Switching.

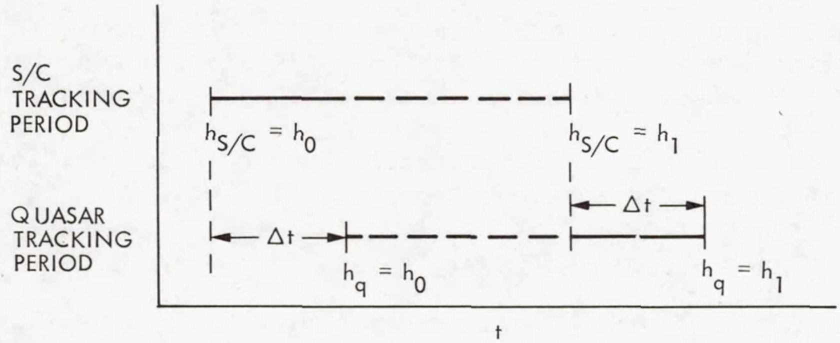


Figure 10. Narrowband ΔVLBI Tracking Scenario

REFERENCES

1. Renzetti, N.A., R. B. Miller and A. J. Siegmenth, "Communications at Planetary Distances," Accademia Nazionale Dei Lincei, Atti Dei Convegni Lincei, 24, Convengo Internazionale, Celebrazione Nazionale Del Centenario, Della Nascita Di Guglielmo Marconi, 1974.

2. Melbourne, W. G., "Navigation Between the Planets," Scientific American, June 1976, Vol. 234, No. 6.

3. Hamilton, T.W., and W.G. Melbourne, "Information Content of a Single Pass of Doppler Data from a Distant Spacecraft," Jet Propulsion Laboratory, Space Program Summary Number 37-39, Vol. III, May 31, 1966.

4. Curkendall, D.W., and S.R. McReynolds, "A Simplified Approach for Determining the Information Content of Radio Tracking Data," Journal of Spacecraft and Rockets, 6:520-525, May 1969.

5. Christensen, C.S., and H.L. Siegel, "On Achieving Sufficient Dual Station Range Accuracy for Deep Space Navigation at Zero Declination," paper presented at the 1977 AAS/AIAA Conference, Jackson Hole, Wyoming, September 7-9, 1977.

6. Jordan, J.F., "Future Challenges in Space Navigation," presentation made at the Institute of Navigation, National Aerospace Meeting, Denver, Colorado, April 13, 1977.

7. Bare, C.C., B.G. Clark, K.I. Kellerman, M.H. Cohen, and D.L. Jauncey, "Interferometer Experiment with Independent Local Oscillators," Science, Vol. 157, pp. 189-191, July 1967.

8. Williams, J.G., "Very Long Baseline Interferometry and Its Sensitivity to Geophysical and Astronomical Effects," DSN Space Programs Summary 37-62, Vol. II, pp. 49-55. Jet Propulsion Laboratory, Pasadena, Calif., March 31, 1970.

9. Rogers, A.E.E., "Very Long Baseline Interferometry with Large Effective Bandwidth for Phase-Delay Measurements," Radio Sci., Vol. 5, No. 10, pp. 1239-1247, October 1970.

10. Hinteregger, H.F., I.I. Shapiro, D.S. Robertson, G.A. Knight, R.A. Ergas, A.R. Whitney, A.E.E. Rogers, J.M. Moran, T.A. Clark, and B. F. Burke, "Precision Geodesy via Radio Interferometry," Science, 178, 396-398, 1972.

REFERENCES (cont.)

11. Thomas, J.B., "An Analysis of Long Baseline Radio Interferometry," DSN Technical Report 32-1526, Vol. VII, pp. 37-50. Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1972.

12. Thomas, J.B., "An Analysis of Long Baseline Radio Interferometry, Part II," DSN Technical Report 32-1526, Vol. VIII, Jet Propulsion Laboratory, Pasadena, Calif., pp. 29-38, April 15, 1972.

13. Thomas, J.B., "An Analysis of Long Baseline Radio Interferometry, Part III, DSN Technical Report 32-1526, Vol. XVI, Jet Propulsion Laboratory, Pasadena, Calif., pp. 47-64, June 15, 1973.

14. Thomas, J.B., J.L. Fanselow, P.F. MacDoran, J.J. Spitzmesser, and L.J. Skjerve, "Radio Interferometry Measurements of a 16-km Baseline with 4-cm Precision," DSN Technical Report 32-1526, Vol. XIX, pp. 36-54, Jet Propulsion Laboratory, Pasadena, CA February 15, 1974.

15. Ong, K.M., P.F. MacDoran, J.B. Thomas, H.F. Fliegel, L.J. Skjerve, D.J. Spitzmesser, P.D. Batelaan, S.T. Pain, and M.G. Newsted, "A Demonstration of Radio Interferometric Surveying Using DSS 14 and the Project Aries Transportable Antenna," DSN Progress Report 42-26, pp. 41-53, Jet Propulsion Laboratory, Pasadena, CA, April 15, 1975.

16. Rogers, A.E.E., "A Receiver Phase and Group Delay Calibration System for Use in VLBI," Haystack Observatory Technical Report 1975-6, Feb. 15, 1976.

17. Thomas, J.B., "The Tone Generator and Phase Calibration in VLBI Measurements," DSN Progress Report 42-44, Jet Propulsion Laboratory, Pasadena, Calif., pp. 63-74, April 15, 1978.

18. Cohen, E.J., J.L. Fanselow, G.H. Purcell, Jr., D.H. Rogstad, and J.B. Thomas, "Recent Astrometric Observations on Intercontinental Baselines," (Advanced VLBI System Group, Jet Propulsion Lab., Pasadena, California). National Radio Sci. Meeting, Univ. of Colorado, Boulder, Colorado, Jan 9-13, 1978.

19. Curkendall, D.W., "Radio Metric Technology for Deep Space Navigation: A Development Overview," Pre-print 78-1395, AIAA/AAS Astrodynamics Conference, Palo Alto, CA, August 7-9, 1978.

20. Melbourne, W.G., D.W. Curkendall, "Radio Metric Direction Finding: A New Approach to Deep Space Navigation," presented at AAS/AIAA Astrodynamics Specialists Conference, Jackson Hole, Wyoming, September 7-9, 1977.

21. Preston, R., et al., "Establishing a Celestial VLBI Reference Frame: I. Searching for VLBI Sources," DSN Progress Report 42-46, Jet Propulsion Laboratory, Pasadena, CA, June 21, 1978.

Questions and Answers

DR. TOM CLARK, NASA Goddard Space Flight Center:

I might update a couple of comments that were made in that one. I believe there has been one successful spacecraft-to-quasar delta VLBI experiment which was done by Shapiro et al. at MIT, and Newhall et al. in your shop, on the differential position of the Viking Orbiter, and the OJ-287 quasar. That was done about a year ago, and the data is still being processed, but it appears that it was successful.

I just wanted to stress one thing in terms of the angular measures that were mentioned here. The units, milliseconds of arc and so forth, were stressed a few times, and I just wanted to remind you what a millisecond of arc was.

If I took this quarter and gave it to Dave, and told him to take it back to Pasadena and tried to look at it from here, that is about a millisecond of arc. That is half a nanoradian.

I would point out that there is at least one pair of quasars which have been done to a few tens of picoradians in terms of differential positions by a delta VLBI type technique. They are only a degree apart on the sky, but it is the size, roughly, of George's eyeball on this thing, in the same analogy as before.

MR. CURKENDALL:

Yes. Not only has there been one experiment with the Viking spacecraft, there are something like 20, an unknown number of which are successful, sitting in our data hoppers.

By my remark, I meant I don't think anybody has been ever able to run an experiment where the accuracy had to be seven centimeters, and then be able to prove that it was. And I think that is accurate; it has probably never been done. And I think this is the same thing whether you can really wring that kind of performance out of the system or not.

DR. CARROLL ALLEY, University of Maryland:

Unfortunately, I arrived late for your talk. You may have mentioned this at the beginning. But it is worth pointing out, I think, that with this coherent tracking, and with a good transponder on these deep spacecraft, that there is a potential of measuring low frequency gravitational waves that may well be occurring in the universe. That is, the part in 10^{14} that you are striving for is about three orders of magnitude too insensitive, according to the current estimates of my friend Kip Thorne and other people.

Nevertheless, I would like to submit, as in an earlier discussion today, that we do not know everthing about the universe, and that there may well be gravitational radiation of higher amplitudes

than is predicted, and I hope that you will be keeping a very careful watch, even at the part in 10^{14} level. The strain induced, the delta L over L in the distance between the spacecraft and the earth, is the same as the strength of the graviational radiation.

MR. CURKENDALL:

I understand.

DR. ALLEY:

And so there may just be some relic radiation of this amplitude around that would show up in such measurements.

MR. CURKENDALL:

Yes. I think your word "relic" is the key here. The thing I don't like about the Kip Thorne calculations is it goes something like this: If you assume the collapse of something like 10^8 solar masses, you will get a differential movement between the spacecraft and the earth.

And maybe this occurs something like once every 30 years. You will get a differential change in length between the earth and the spacecraft of about one and a half millimeters; and the problem of sitting around for 20 years waiting for that to happen is immense.

What seems more feasible to me is to look at what you have just said: Sit there and look for the background radiation. If you can put an X-band uplink on the spacecraft, and get masers down to a few parts in 10^{15}, the numbers already work out that you should be able to see enough gravitational energy that would close the universe. And that is much better than waiting around.

And the advantage of that experiment is if you go out and you look and you don't see anything, you can go home. And that is really important.

DR. VICTOR REINHARDT, NASA Goddard Space Flight Center:

Just one question and it is a partial comment: Is there any possibility of using telemetry information to increase the bandwidth and narrowband VLBI to give you an effective wideband?

MR. CURKENDALL:

The telemetry, right now, runs about 360 kilohertz subcarriers and at around 100 kilobits a second. You see, relative to the earth's satellites, where the signal and the noise are so much better, the telemetry rate coming back is not that high. And that is not very much bandwidth spreading.

We are doing that as part of this demonstration program. In fact, we have a mode where we leave the subcarrier on, but turn off the modulation so you get the nice pure sine waves, and you look at the harmonics of those sine waves.

It is just like the Goddard side-tone ranging system, in terms of being able to detect that.

But we detect it through the open loop RF chain, and calibrate with the quasar signals. And so, yes; you can do it as a restricted bandwidth sort of thing.

What we hope to put on Galileo is tones of plus-minus 20 megahertz on down, and that ought to do it.

MR. DAVID W. ALLAN, National Bureau of Standards:

I would think that this level of sensitivity, you would be sensitive to lunar crustal tidal movements of the mantle of the earth. This was not mentioned. I was just wondering if this is a problem. Is this not of the order of a meter or two? I am not sure.

MR. CURKENDALL:

Yes, they are; and I think our Chairman can give better words on that. I think the bottom line is that you have to model it; but it is thought to be modelable.

DR. TOM CLARK, NASA Goddard Space Flight Center:

Yes, there are a number of these little subtle, insidious effects that affect the stations on the earth. The particular one with tides is not as bad as many of the others because the tides have a semi-diurnal signature, and it is very easy to pull semi-diurnal signatures out, because all of the things that Dave was talking about were diurnal signatures. It is those diurnal ones that are much more insidious and give you a lot more trouble; things like diurnal polar motion, for example.

MR. CURKENDALL:

Do you know what you are most sensitive to? I once did a spectrum analysis, and for reasons I have never quite understood, you are really most sensitive to twice diurnal rates. With a given amount of power, you are more sensitive to something with a 12-hour period than a 24-hour period. I have never quite understood that, physically.

Page intentionally left blank

TEMPUS

A PROPOSAL FOR AN INTERNATIONAL TIME TRANSFER AND PRECISION TRACKING SATELLITE

David C. Holmes

Since W. G. Cady carried his piezo resonator to seven laboratories in England, France, Italy and the United States in 1923, the primary method of coordinating time between international stations has involved the use of traveling clocks. These were first transported by ship and train, more recently by aircraft. Project TEMPUS also requires traveling clocks but these will revolve thousands of kilometers above the earth in artificial satellites.

Several methods of international time transfer are now available and a comparison of various concepts derived from data published by the European Satellite Agency is shown below.

Method	Accuracy	Remarks
VLBI (pulsars)	1 nsec	Slow, expensive ground stations.
TV relay via satellite	10 nsec	Requires wide-band satellite transponder with satellite visible from all stations.
Portable clocks	30 nsec	Slow, spasmodic, expensive.
NTS	200 nsec	Better clocks and orbits are needed.
Loran-C	200 nsec	Limited by propagation phenomena.

In addition to these systems, ESA is scheduled to launch the SIRIO 1 in 1979 and the LASSO/SIRIO 2 experiment in 1981. SIRIO 1 will

employ microwave transponder time synchronization and LASSO/SIR1O 2 will combine both microwave and optical tracking capabilities since it will carry laser retroreflectors. Basically, the system will allow for the comparison of a ground station timing signal with that of a satellite time reference clock. It will be an excellent demonstration of satellite capability to synchronize time and at the same time furnish a source of data for the study of atmospheric path time delay errors in the electromagnetic frequency spectrum. The SIR1O satellites are scheduled to be placed in geostationary orbits.

Satellite experiments to provide time synchronization began with the launch of Timation II in 1969 by the U.S. Naval Research Laboratory. The Timation series were the forerunners of the present Navigation Technology Satellites (NTS) which NRL is providing to support the NAVSTAR Global Positioning System Program. Figure (1) illustrates the method used. The satellite clock is updated by time from the U.S. Naval Observatory and is used to correct the time of remote clocks. The accuracy of time closure is sensitive to errors caused by unknowns in the clock drift and satellite position (ephemeris) as well as refractive errors in the atomospheric transmission path.

By the summer of 1978 there was considerable world-wide interest in time transfer by means of satellites, and stations in England, France, Germany, Spain, Japan, Australia and Canada as well as those in the U.S. were working with the NTS spacecrafts. Figure (2) shows the extensive network involved.

During 1978 an experiment was conducted to compare the results of time transfer via portable clock with that achieved by the use of NTS time receivers. As can be seen from figure (3), the agreement in all cases was within a microsecond.

In 1973, the NAVSTAR Global Positioning System (GPS) program was established by the U.S. Department of Defense, and it combined NRL's Timation with the U.S. Air Forces' 621B satellite navigation project. This action had a considerable impact not only on the development of navigation satellites, but also on the future of satellite time transfer.

Solution of the navigation equation by the use of satellites is inextricably tied to satellite time transfer since both require the same fundamental elements; excellent clocks with highly predictable drift and accurate knowledge of the satellite ephemeris. In addition, if reliable continuously available world-wide time transfer is to be accomplished from space, a multiple satellite system will be required for it is only by the use of a number of

orbiting satellites, each having excellent clocks and well defined orbits, that a dependable on-demand time transfer system can be developed.

Similar characteristics are required for a satellite which can be used in conjunction with either lasers or electromagnetic receivers for precision tracking purposes such as the determination of earth plate movements and the refinement of our knowledge of the relativity and gravitational constants.

Perhaps the most obvious and practical satellite candidates for such a system are a selected and modified version of a number of the NAVSTAR GPS spacecraft which are now in development. These will become operational within the next decade. It is not necessary that all of these spacecraft be configured as TEMPUS satellites. The full NAVSTAR system will contain 24 birds arranged in three planes, each at an inclination of 45 degrees. They will be spaced in such a manner that at least four satellites will be in view anywhere on the globe at all times. Each will carry an updated ephemeris and four or more will provide users with the elements of a three dimensional navigation fix plus a time update. The full NAVSTAR constellation is shown in figure (4). The system will orbit the earth at 17,300 kilometers, easily within laser range if adequate retroreflectors are carried.

The three dimensional navigation accuracy available from the NAVSTAR satellites will be extremely high by today's standards. Initial tests at Yuma, Arizona, conducted under controlled conditions, have reported position fixing to better than 10 meters. However, this accuracy is probably not good enough for many precise scientific research projects in geodynamics and relativity theory. For these and similar purposes accuracies in the centimeter range are necessary.

For this reason it has been interesting to design a TEMPUS package which, when added to a satellite similar to NAVSTAR would provide a precision scientific tool capable of meeting the stringent requirement of the scientific community, while at the same time not interfering with the basic GPS operational mission. This package should be designed to have minimum impact upon the basic NAVSTAR configuration and should be able to function normally in the GPS operational constellation. An initial estimate of the number required to provide worldwide time synchronization as well as a precise tool for research experiments needing such a capability is two TEMPUS satellites in each plane for a total of six. Properly spaced, this number should provide a constellation of four in view for almost all parts of the world on a periodic basis. Properly configured the TEMPUS package will define position

accuracies in the centimeter range for stationary locations on the earth's surface in addition to international time synchronization to about one nanosecond.

In order to provide the frequency stability necessary for such accuracies, the TEMPUS satellite must take advantage of a navigation package similar to that already avilable in the NAVSTAR system. Using the NAVSTAR principle, the clock in the satellite is synchronized with that of the user. The satellites signal is time coded thus allowing accurate measurement of the distance from satellite to observer. In the case of a station whose position is known, the signal from one satellite only is required to effect time transfer.

It can be readily seen that the heart of such a system is an accurate and very stable clock. The first clocks flown in Timations I and II were quartz and had stabilities of about 10^{-11} parts per day. NTS-1 contained a rubidium atomic clock whose stability was about 10^{-12} parts per day while NTS-2, launched in June 1976, had a cesium standard whose stability is on the order of 10^{-13} NTS-3 which is scheduled for flight in 1982 and the operational NAVSTAR system are designed to contain advanced cesium or hydrogen maser clocks whose stability should approach 10^{-14} per day. Accuracy of this order is required to provide international nanosecond time synchronization and positioning in the centimeter range. Figure (5) is a diagram of NTS-1 and figure (6) shows NTS-2. The configuration of NTS-3 will be quite similar.

To provide the capabilities desired for the TEMPUS satellites, more than the NAVSTAR navigation package and an accurate clock is needed. For additional precise tracking, laser retro-reflectors are desirable. These were carried on NTS-1 and again on NTS-2 which has 44 corner reflectors on the bottom of the satellite. These were successfully tracked with the Smithsonian Astrophysical Observatory's laser on Mt. Hopkins, Arizona. Figure (7) shows a picture of the laser corner reflectors and figure (8) indicates the results of the Mt. Hopkins laser tracking experiment which produced errors whose sigma was less than two meters in this first attempt.

In addition to laser retroreflectors, the TEMPUS package should contain accelerometers to measure external forces on the satellite such as solar pressure since this is probably the largest single source of ephemeris error. An event timer should also be added to record significant happenings aboard the spacecraft. Because of the development of very long baseline interferometry receiver capability, and the measurement precision available with this tool,

the TEMPUS package should also provide a signal for VLBI reception.

Since the navigation package will already be aboard the NAVSTAR satellites, the modified TEMPUS spacecrafts should contain at least the following extra units:

- o Hydrogen maser/advanced cesium clock,
- o Laser retroreflectors,
- o Accelerometers,
- o Event timer,
- o VLBI transmitter.

A relatively simple TEMPUS parkage can be designed to provide all these components at relatively small cost in weight, space and power. In addition to providing much greater accuracy this package will also allow improved resolution of the ephemeris error budget components and thus a better ephemeris prediction model. The NTS-3 satellite can provide all these capabilities. Since this satellite is scheduled for launch in 1982, it will be an excellent prototype TEMPUS satellite, particularly since it is to act as a replenishment spacecraft in the NAVSTAR constellation.

Cost and timescale information for TEMPUS will depend upon requirements and future world economic conditions, however the basic technology is currently available and will be tested in NTS-3. By the 1982 timeframe the space shuttle should greatly reduce launch costs.

It should be emphasized that the development of a time synchronization and precision tracking satellite package such as TEMPUS for installation aboard selected NAVSTAR satellites or some other system with similar characteristics should be a truly international program with the participation of many countries if we are to derive maximum scientific research benefits from its use.

If the need for such a capability in space is visualized for the period beginning in 1983, now is the time to begin the planning. It is suggested that those scientific groups who foresee a requirement for such a precise time synchronization and tracking system take the lead in establishing the system parameters.

NAVSTAR GPS
NAVIGATION TECHNOLOGY SEGMENT

STATION SYNCHRONIZATION
BY
TIME TRANSFER

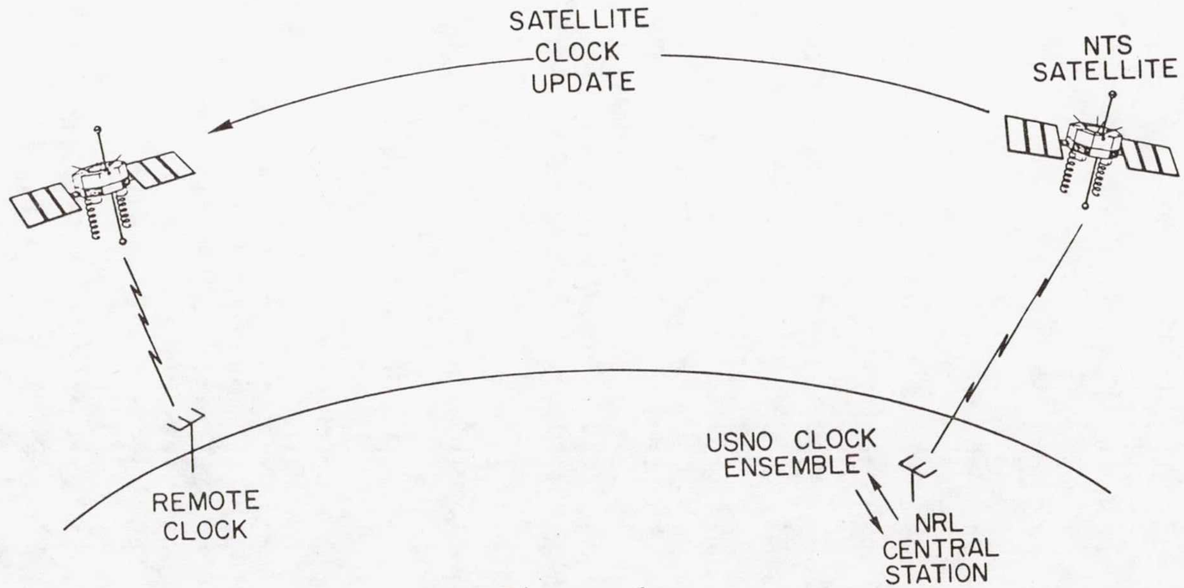


Figure 1

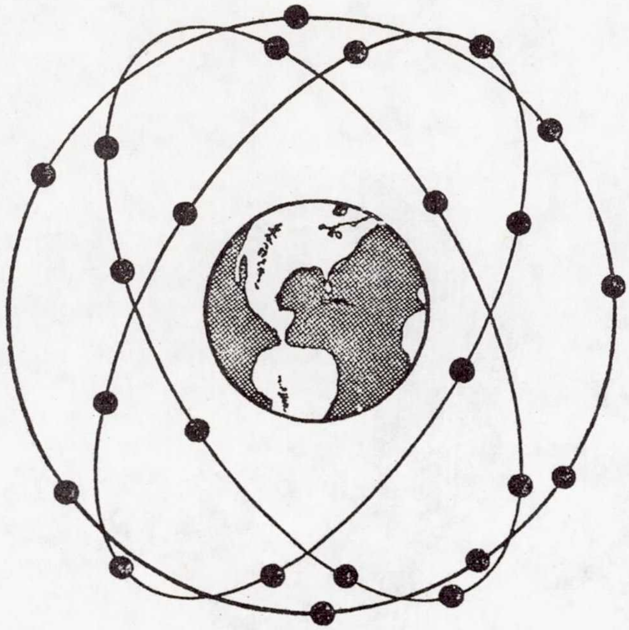
Figure 2

INTERNATIONAL TIME SYNCHRONIZATION VIA NAVIGATION TECHNOLOGY SATELLITE (ITS / NTS)

SUMMARY OF
PORTABLE CLOCK CLOSURES
VS
NTS TIME TRANSFER RESULTS

STATION	DAY (1978)	PORTABLE CLOCK NTS TIME TRANSFER (US)
BIH	124	-.57
CERGA	117	.70
DNM	282	.09
IFAG	199	.03
NBS	221	.19
NRLM	299	-.53
RGO	115	.44
RRL	303	.13
USNO	186	.04

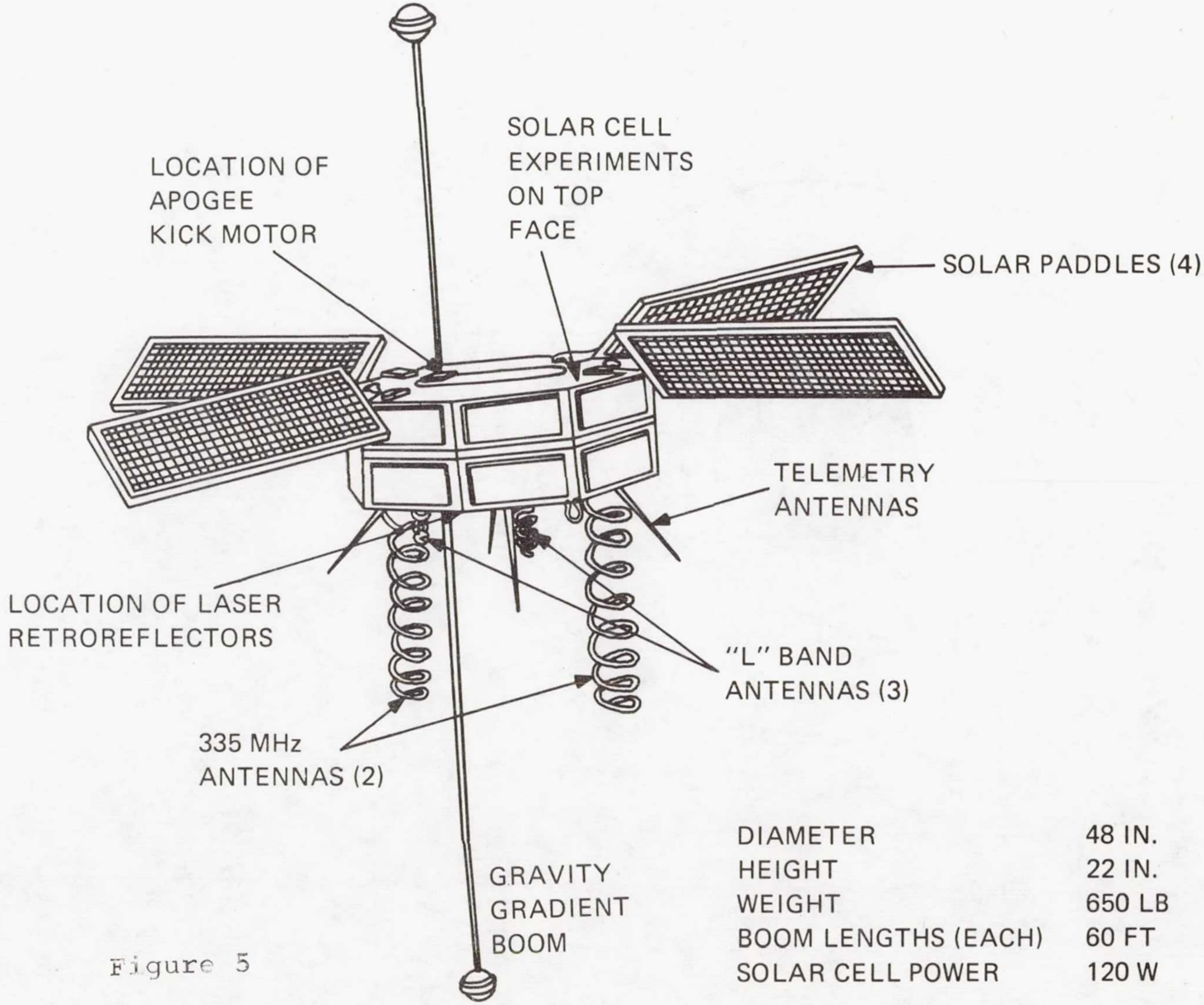
Figure 3



FULL CAPABILITY
24 SATELLITES (MID 1980's)

Figure 4

NTS-1



DIAMETER	48 IN.
HEIGHT	22 IN.
WEIGHT	650 LB
BOOM LENGTHS (EACH)	60 FT
SOLAR CELL POWER	120 W

Figure 5

NTS-2

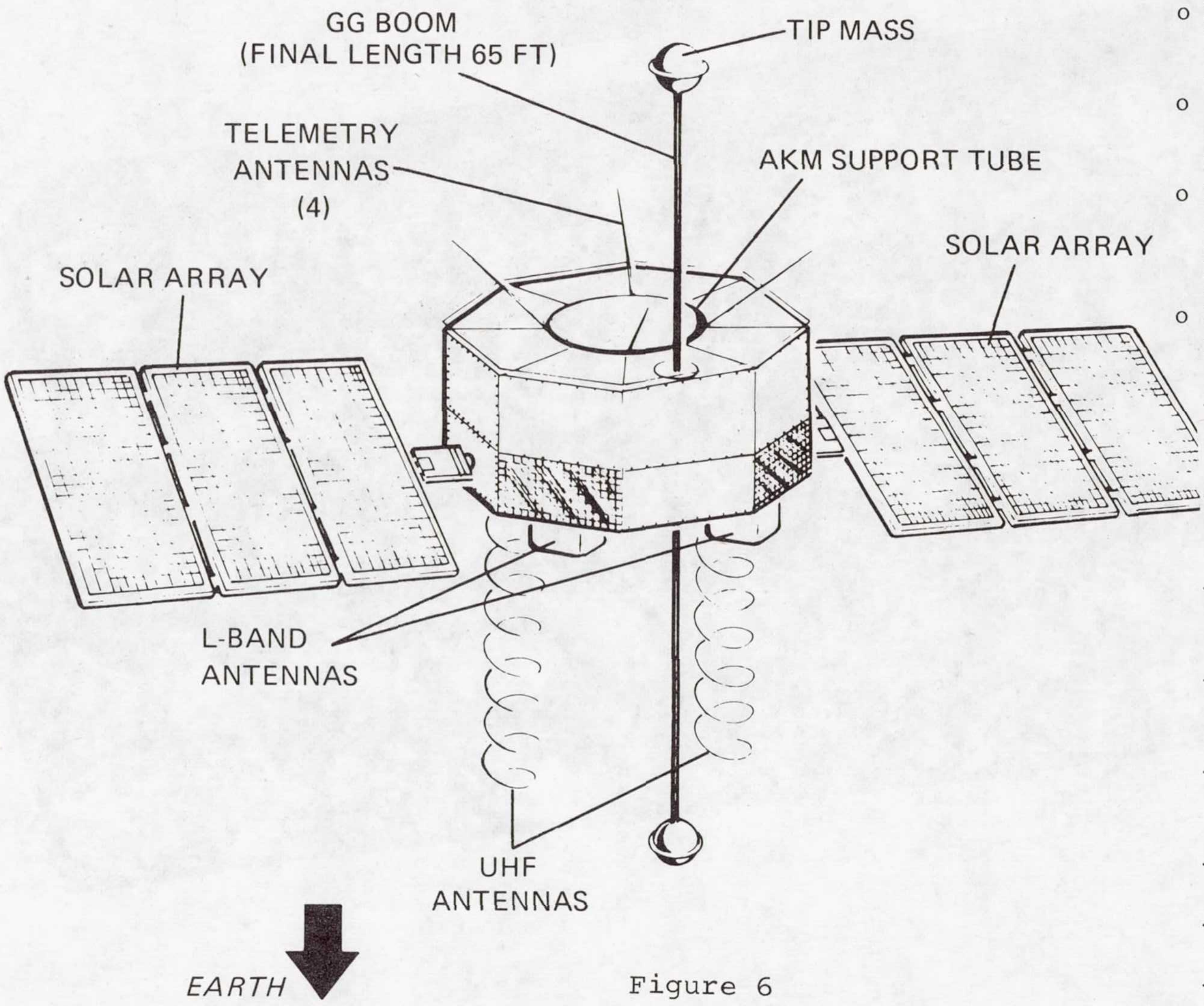


Figure 6

OBJECTIVES

- o Demonstrate Feasibility of Cesium Frequency Standard
- o Initial Demonstration of Navigation Payload
- o Function as One of Satellites in GPS-Phase I Constellation
- o Specific Technology Objectives:

  - -To Provide Orbit Prediction & Tracking Concepts
  - -Test the Use of Nickel-Hydrogen Batteries
  - -Test Side-Tone Ranging Navigation
  - -Time Synchronization Between Remote Stations
  - -Stage/Launch Vehicle Performance
  - -Laser Retroreflector
  - -Solar Cell Experiments
  - -Thermal Coating Experiment

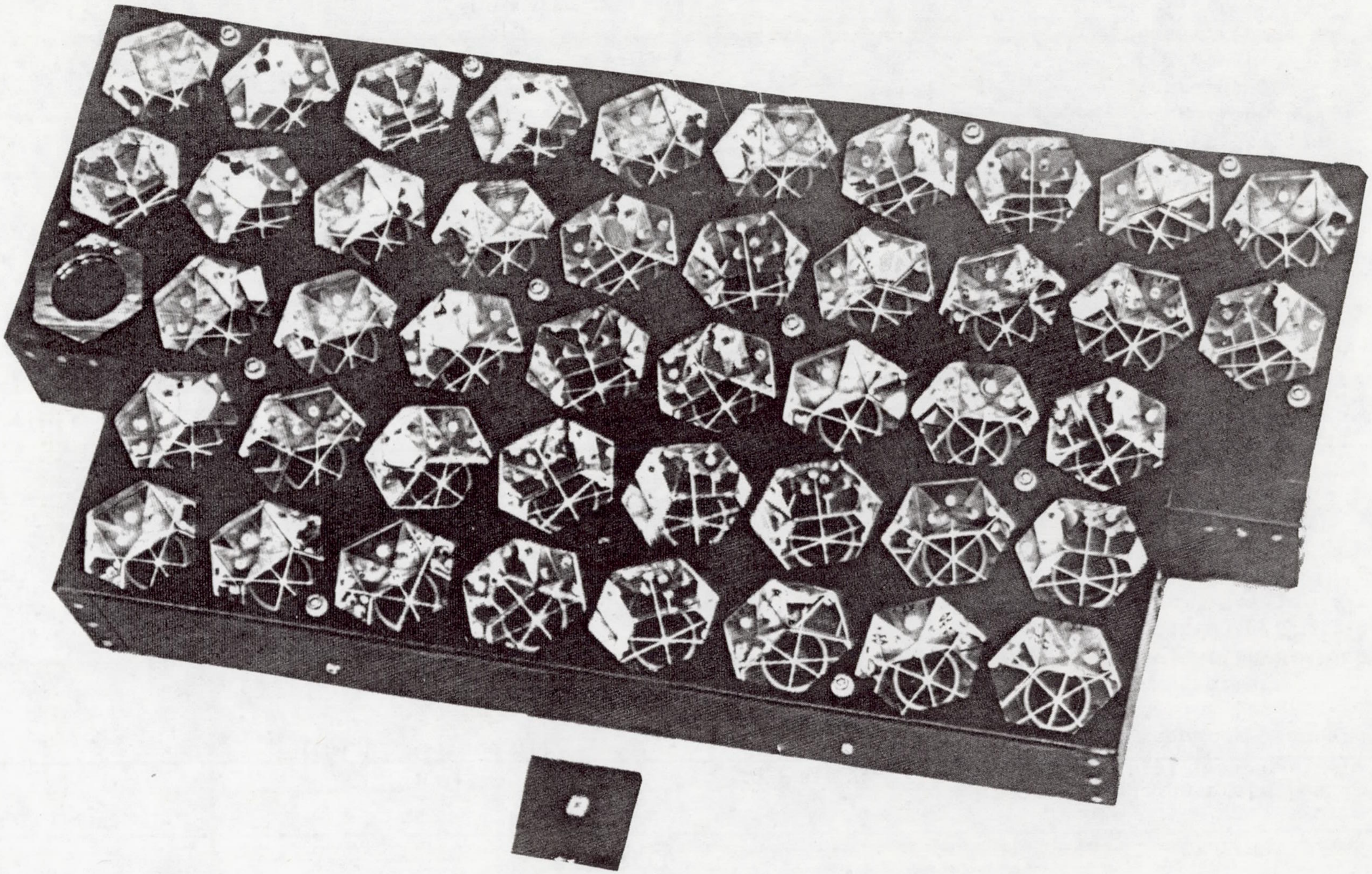


Figure 7
Laser Retroreflector NTS-2

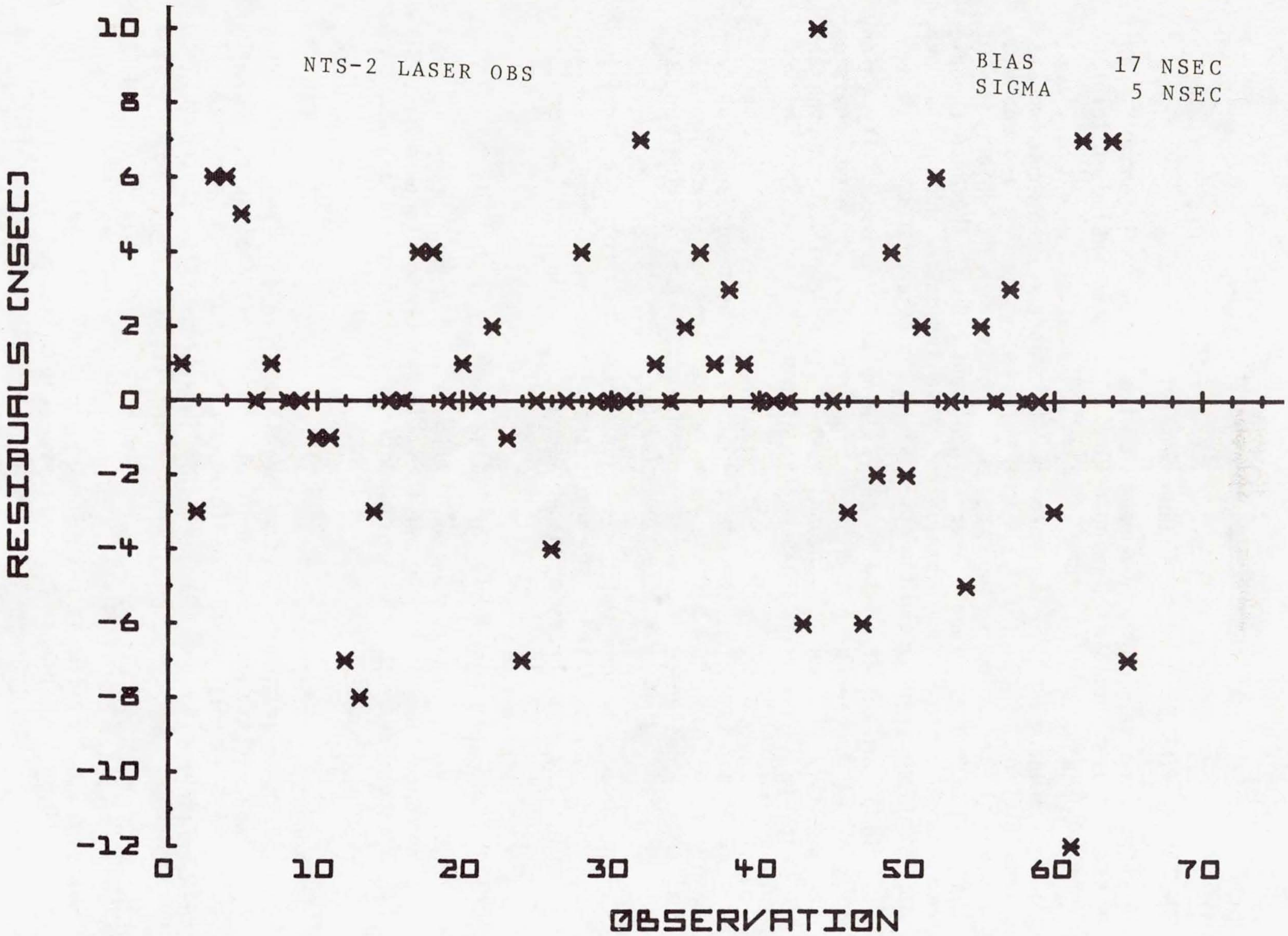

NTS-2 LASER OBS
BIAS 17 NSEC
SIGMA 5 NSEC
RESIDUALS (NSEC)
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
0
10
20
30
40
50
60
70
OBSERVATION

FIGURE 8

QUESTIONS AND ANSWERS

DR. GERNOT M. R. WINKLER, U.S. Naval Observatory:

I think I am very much in agreement with the overall concept. However, I think that your timing estimate is conservative at least by a factor of 10.

Laser ranging, two-way ranging, if it is executed from two stations within short intervals, does not really require a great stability of the clock; and the experiments at the University of Maryland and of the lunar laser ranging and others, the NASA experiments with the geodetic satellites, have indicated that nanosecond and fractional nanosecond time precisions can be expected.

So I think you probably are talking about a capability which will be as far-reaching and of as high precision as we can envision can possibly be used by anyone, and will be available during the foreseeable future. And this makes, of course, such a proposal very interesting.

I do not feel that the constellation may be necessary, but certainly more than one, because we have seen from past experience the fallacy of designing expensive experiments around a single satellite. I think that is a very fallacious approach.

However, the one which you have recommended is, of course, going around that. I think the whole thing makes very much sense, and I would like to express my hope that such a thing can be promoted. Thank you.

DR. HELMUT HELLWIG, National Bureau of Standards:

I would like to add a question. As most of you know, there is also a European Space Agency proposal on a related package called the LASSO experiment. Is there international coordination between that and TEMPUS?

MR. HOLMES:

I think there are probably other people who probably know more about that experiment than I do. My understanding is that it is a single shot which is piggyback, I believe, aboard an Italian satellite. It has a useful lifetime of on the order of five months; is that right, Dr. Winkler?

DR. WINKLER:

No. The useful life is much longer. However, the availability for the Western Hemisphere is only five months because it will be rescheduled to go somewhere around longitude zero, or even +15°, and it will not be available to us any more. And in addition, it is a

very expensive one to instrument and it is a single one. And the objections and questions I have indicated before apply here very much.

DR. HELLWIG:

It is on a synchronous satellite. But as I understand the European people, although what you say is true, they plan for the future to put the package on other vehicles which are not necessarily just for the European sphere. It remains to be seen.

Page intentionally left blank

10th PTTI REGISTRATION LIST

Allan, David W.
National Bureau of Standards
Boulder, CO 80303

Allen, Jerry J.
USNWC
China Lake, CA 93555

Allen, Ralph T.
Naval Electronic Systems Command
Code 51023
Washington, DC 20360

Alley, Carroll
University of Maryland
College Park, MD 20742

Alpert, Arnold
USAF, Newark AFS
Newark, OH 43055

Anderle, Richard J.
Naval Surface Weapons Center
Dahlgren, VA 22448

Anderson, Charles R.
Hughes Aircraft Co.
P.O. Box N
Aurora, CO 80041

Badura, Heinz
Efratom California, Inc.
18851 Bardeen Avenue
Irvine, CA 92715

Bates, Alvin
Johns Hopkins University
Johns Hopkins Road
Laurel, MD 20810

Baker, Robert E.
U.S. Department of Defense, R921
9800 Savage Road
Ft. Meade, MD 20755

Bartholomew, C. A.
NRL, Code 7962
Washington, DC 20375

Beaty, James S.
FAA NAFEC ANA 180
Atlantic City, NJ 08405

Beard, Ronald
NRL
Washington, DC 20375

Besson, Raymond J.
E.N.S.C.M.B. Besanson, France/N.B.S.
325 Broadway
Boulder, CO 80303

Bloch, Martin B.
FEI
3 Delaware Drive
New Hyde Park, DE 11040

Blomberg, Eric
Smithsonian
60 Garden Street, P151
Cambridge, MA 02138

Bowerman, Jack C.
NAVELEX
NAVELEX 46014
Washington, DC 20360

Bregstone, Ed
U.S. Coast Guard
400 Seventh Street
Washington, DC 20591

Brown, Joe A.
NWC China Lake, CA
400 E. California Blvd.
Ridgecrest, CA 93555

Bryant, Ellis H., Jr.
The Weatherchron Co.
4881 Powers Ferry Road
Atlanta, GA 30327

Buisson, James
NRL, Code 7960
Washington, DC 20375

Burke, Guy W.
Naval Electronic Systems Command
Arlington, VA 22020

Busca, Giovanni
Ebauches Sa
Dir Recheches
Ch 2001 Neuchatel
Switzerland 038247960

Bush, Robert L.
Bendix
9250 Route 108
Columbia, MD 21045

Carter, William E.
National Ocean Survey
Rockville, MD 20852

Cashion, Roy
Satellite Navigation
2141 Industrial Parkway
Silver Spring, MD 20904

Cervenka, Dr. Peter
Phoenix Corporation
P.O. Box 2797
Laurel, MD 20811

Chi, Andrew R.
NASA-GSFC
Code 810
Greenbelt, MD 20771

Clark, Dr. T. A.
NASA/Goddard
Code 693
Greenbelt, MD 20771

Clarke, Randolph T.
USNO
Washington, DC 20390

Cloeren, James M.
Satellite Navigation
2141 Industrial Parkway
Silver Spring, MD 20904

Coffin, Richard C.
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91214

Costain, Dr. Cecil
National Research Council of Canada
Division of Physics, NRC
Ottawa, Canada K1A0R6

Cox, Dr. Duncan B., Jr.
C. S. Draper Lab., Inc.
MS 92, 555 Tech. Sq.
Cambridge, MA 02139

Cresswell, Stephen A.
Federal Electric
VAFB, CA 93436

Curkendall, David
JPL
4800 Oak Grove Drive
264–764
Pasadena, CA 91214

Danzy, Frederick
NRL
4555 Overlook Avenue
Washington, DC 20375

Davis, Duane G.
Hughes Aircraft Co.
P.O. Box N
Aurora, CO 80041

Decker, B. Louis
DMA Aerospace Ctr.
402 Devon Court
Ballwin, MO 63011

Derengowski, Walter L.
T. E. Corp.
3702 Perry Avenue
Kensington, MD 20795

Donaldson, R. W., Jr.
Westinghouse Elec.
P.O. Box 1897, MS 679
Baltimore, MD 21203

Doepke, Lawrence
DMAAC GDDT
St. Louis AFS, MI 63118

Ealum, Richard L.
DMAAC GDDT
St. Louis AFS, MI 63118

Easton, Roger L.
NRL, Code 7960
Washington, DC 20375

Echols, Jim
Austron, Inc.
1915 Kramer Lane
Austin, TX 78758

Edwall, Warren E.
RCA Corp.
1901 N Moore St.
Arlington, VA 22209

Ellis, Bob
Austron, Inc.
1915 Kramer Lane
Austin, TX 78758

English, Dr. Thomas C.
Efratom Systems Corp.
18851 Bardeen Avenue
Irvine, CA 92715

Emmons, Don
Frequency & Time Sys.
182 Conant Street
Danvers, MA 01923

Erb, Lee A.
Erbtec
3565 Catalpa Way
Boulder, CO 80302

Fabrizio, William
General Dynamics
10221 Grayfox Drive
San Diego, CA 92131

Finsod, Theodore D.
Navy Astronautics
Point Mugu, CA 93042

Fischer, Mike
Hewlett-Packard
5301 Stevens Creek Blvd.
Santa Clara, CA 95050

Fisher, Laura C.
U.S. Naval Observatory
34th & Mass. Ave., N.W.
Washington, DC 20390

Fleming, Robert F.
Satellite Navigation
2141 Industrial Parkway
Silver Spring, MD 20904

Folen, Vincent J.
NRL
Code 5292
Washington, DC 20021

Frensch, Theo
Hewlett-Packard
5301 Stevens Creek
Santa Clara, CA 95050

Fruehauf, Hugo
Efratom
18851 Bardeen Avenue
Irvine, CA 92715

Fulper, Richard, Jr.
NRL
8006 Carey Branch Drive
Oxon Hill, MD 20022

Galloway, Larry M.
USAF-AGMC
Newark AFS
Newark, OH 43055

Georgevic, Dr. Robert R. M.
University of Petroleum and Minerals and
General Dynamics/Electr.
2325 Woodlyn Road
Pasadena, CA 91104

Gouin, Michel
Radio-Canada
1400 Dorchester E.
Montreal, Quebec, Canada H2L2M2

Gubser, Donald
NRL
Washington, DC 20375

Graden, Maynard, A., III
Siemens Corporation
5500 Bronen Snd. Blvd.
Boca Raton, FL 33431

Grant, Athay R.
NCAR
P.O. Box 3000
Boulder, CO 80307

Halford, Donald
NBS 723ASM
Boulder, CO 80303

Hall, R. Glenn
U.S. Naval Observatory
Washington, DC 20390

Hanson, Wayne
National Bureau of Standards
325 South Broadway
Boulder, CO 80302

Harmatuk, Samuel N.
Consultant
1575 Odell Street
Bronx, NY 10462

Hellwig, Helmut
National Bureau of Standards
Boulder, CO 80303

Henry, C. Patrick
Frequency Elec.
3 Delaware Drive
New Hyde Park, NY 11040

Hesselberth, R. J.
Spectracom Corp.
1667 Penfield Road
Rochester, NY 14625

Hiraiwa, Dave S.
IEEE-IM
P.O. Box 22024
Honolulu, HI 96822

Ho, John
Frequency Electronics
3 Delaware Drive
New Hyde Park, NY 11040

Howatt, J. R.
C. S. Draper Labs
555 Technology Square
Cambridge, MA 02154

Howell, Delmos D.
Western Electric Co.
4513 Western Avenue
Lisle, IL 60532

Hoen, John
R942
Dept. of Defense
9800 Savage Road
Fort Meade, MD 20755

Holmes, David C.
NRL, Code 7960
612 7th Street
Washington, DC 20375

Hruska, Carl Karel
York University
4700 Keele Street
Downsview, Ontario, Canada M3J1P3

Huebner, Dr. Udo
Phys. Techn. Bundesanstalt
Bundesallee 100
Braunschweig, Germany

Hurd, William J.
Jet Propulsion Lab
4800 Oak Grove Drive
Pasadena, CA 91103

Hynes, John F.
U.S. Air Force
HQ1842EEG AFCS
Scott AFB, IL 62225

Imbier, Edward A.
Smithsonian Inst.
60 Garden Street
Cambridge, MA 02138

Jansen, Norman J.
USAF
HQ AGME
Newark, OH 43055

Jechart, Ernst
Efratom California
18851 Bardeen Avenue
Irvine, CA 92715

Jeffries, Marie M.
NRL, Code 7966
4555 Overlook Avenue
Washington, DC 20375

Johnson, Andrew C.
U.S. Naval Observatory
Washington, DC 20390

Kahan, Alfred
RADC/ES
Hanscom AFB
Bedford, MA 01731

Kellogg, William W.
Lockheed Missile & Space
Building 154
P.O. Box 504
Sunnyvale, CA 94088

Kern, Robert H.
Frequency & Time
182 Conant Street
Danvers, MA 01923

Kimsey, David B.
Harris Corp.
P.O. Box 37
Melbourne, FL 32901

Klepczynski, W. J.
U.S. Naval Observatory
Washington, DC 20390

Kolde, Frank K.
Rockwell International
3370 Miraloma Avenue
Anaheim, CA 92803

Kovacs, Robert M.
Kentron Int., Inc.
P.O. Box 1207, A.P.O.
San Francisco, CA 96555

Kuhnle, Paul F.
Jet Propulsion Lab
4800 Oak Grove Drive
Pasadena, CA 91206

Lasley, Eric
General Dynamics
8635 Hydra Lane
San Diego, CA 92126

Lawrence, Dr. Robert
Vitro Labs
14000 Georgia Avenue
Silver Spring, MD 20910

Lazarus, Felix
Hewlett-Packard AG
19 Ch. Chateau-Bloc
1219 Le Lignon
Geneva, Switzerland

Lee, John H.
USAF
343CRS/MACP/EPTRS
Elmendorf, AK 99506

Lerner, David
350 Bleecker Street
New York, NY 10014

Levine, Martin
Frequency & Time Sys.
182 Conant Street
Danvers, MA 01923

Linardi, Horacio
Observatorio-Naval
Av/España-2099
Argentina

Lukac, Carl F.
Naval Observatory
34th & Mass. Ave.
Washington, DC 20390

Lynch, Donald W.
Naval Research Lab
Washington, DC 20375

Ma, Chopo
NASA/Goddard Space Flight Center
Code 922
Greenbelt, MD 20771

Martin, Charles F.
Defense Mapping Agency
USN OBS, Bldg. 56
Washington, DC 20305

Martin, Thomas D.
Electrospace Systems
P.O. Box 1359
Richardson, TX 75080

McCaskill, Thomas
U.S. Naval Research Lab
Code 7966
Washington, DC 20375

Meirs, Marvin
Frequency Elec., Inc.
3 Delaware Drive
New Hyde Park, NY 11040

May, Charles
Johns Hopkins Apl.
Johns Hopkins Road
Laurel, MD 20810

Metcalf, John L.
Tracor, Inc.
101 Guenther Drive
Great Mills, MD 20634

Metcalf, Wendell D.
NESEA
St. Inigoes, MD 20684

Myerson, Herbert
NASA/GSFC, Greenbelt
1203 Dunoon Court
Silver Spring, MD 20903

Mitchell, D. H.
Kentron Int., Inc.
Box 1207, A.P.O.
San Francisco, CA 96555

Mitchell, Ron L.
NEC America, Inc.
7515 Greenville Avenue
Suite 400
Dallas, TX 75231

Monger, Don R.
U.S. Naval Observatory
Richmond FL 33157

Morgan, Edward M.
General Electric
P.O. Box 8555
Philadelphia, PA 19101

Mueller, Ivan I.
Ohio State Univ.
Dept. Geoid Sci.
1958 Neil Avenue
Columbus, OH 43210

Murphy, Jesse K.
Defense Mapping Agency
6500 Brooks Lane
Washington, DC 20019

Murphy, Hugh J.
Rockwell International Corp.
3370 Miraloma Avenue
Anaheim, CA 92803

Murray, James A., Jr.
Naval Research Lab
Code 7552
Washington, DC 20375

McConahy, R. J.
Appl. Phys. Lab/JHU
2717 Woodedge Road
Wheaton, MD 20906

McGuire, M. D.
NASA/GSFC
2015 Serra Avenue
Santa Clara, CA 95050

McWherter, Irving
Vitro Labs
14000 Georgia Avenue
Silver Spring, MD 20901

Narr, Frank J.
Vitro Laboratories
14000 Georgia Avenue
Silver Spring, MD 20910

Nichols, Stephen
Naval Research Lab
Code 7580
Washington, DC 20375

Norton, J. R.
Johns Hopkins/APL
Johns Hopkins Road
Laurel, MD 20810

Nottarp, Klemens J.
Inst. Ang. Geodaesie
Sat. Obs. St. Wettzell
Koetzting, Federal Republic of Germany D-8493

Oaks, Jay
NRL
Washington, DC 20375

Oja, George W.
NAVPRO Pomona
P.O. Box 2505
Pomona, CA 91766

Orr, William S.
CBSI Ltd.
221 S 85th E. Ave.
Tulsa, OK 74112

Osterdock, Terry N.
Hewlett-Packard Co.
5301 Stevens Creek
Santa Clara, CA 95050

O'Sullivan, Eugene J.
MITRE
P.O. Box 208
Bedford, MA 01730

Owolo, David
Western Union
1 Lake Street
Upper Saddle River, NJ 07458

Parello, Tom
Frequency & Time Systems
182 Conant Street
Danvers, MA 01923

Payne, Charles R., Jr.
Navy Astronautics
Pt. Mugu, CA 93042

Peters, H. E.
Sigma Tau Standards Corp.
P.O. Box 1877
Tuscaloosa, AL 35401

Petrie, Robert G.
Jet Propulsion Lab
4800 Oak Grove Drive
Pasadena, CA 91103

Phillips, David H.
NRL
Code 7524
Washington, DC 20375

Phillips, Gail K.
General Dynamics
P.O. Box 81127
San Diego, CA 92138

Pilkington, John D. H.
Royal Greenwich Observatory
Herstmonceux Castle
Hailsham, E. Sussex
BN 27 IRP, England

Pitsenberger, John W.
U.S. Navy Dept.
Washington, DC 20376

Popa, Adrian E.
Hughes Research Labs.
3011 Malibu Canyon Road
Malibu, CA 90265

Porter, Milton R.
Bendix Field Eng.
9250 Route 108
Columbia, MD 21045

Price, George E.
Austron, Inc.
1915 Kramer Lane
Austin, TX 78758

Putkovich, Ken
U.S. Naval Observatory
Washington, DC 20390

Reinhardt, Victor
NASA/GSFC
Code 814
Greenbelt, MD 20771

Robertson, Douglas S.
NGS NOS NOAA
Rockville, MD 20852

Roloff, R. C.
NASA/Bendix Corp.
12430 Winding Lane
Bowie, MD 20715

Rowley, William L.
Western Electric
29W413 Thornwood
Warrenville, IL 60555

Rueger, Lauren J.
Johns Hopkins Univ.
Johns Hopkins Road
Laurel, MD 20810

Ryan, Jim
NASA/GSFC
9627 Carriage Road
Kensington, MD 20795

Schmid, John G.
USAF AFSC SAMTEC
Vandenberg AFB, CA 93437

Scull, David C.
USCG Hqrs. G–ONSOD
2100 2nd Street, S.W.
Washington, DC 20590

Skopetz, Alexander
NASA/GSFC, Code 730.4
Greenbelt Road
Glendale, MD 20715

Smith, Arthur G.
Pacific Missile Test Ctr.
Data Coll. Div., Code 3410–2
Point Mugu, CA 93042

de Souza, Frade V.
Observatorio–Naval–
Av.–España–2099
Argentina

Sridhar, Champa Guha
Hewlett–Packard Co.
5301 Stevens Creek
Santa Clara, CA 95117

Stebbins, Sarah
Code 7966, NRL
4555 Overlook Avenue
Washington, DC 20375

Stein, Samuel R.
National Bureau of Standards
325 Broadway
Boulder CO 80303

Stephens, John E.
Hewlett–Packard Co.
Stevens Creek Blvd.
Santa Clara, CA 95117

Stone, C. S.
Austron, Inc.
1915 Kramer Lane
Austin TX 78758

Stover, Harris A.
DCA/DCEC
1860 Wiehle Avenue
Reston, VA 22090

Strain, John
Frequency Elec.
13975 Conn. Avenue
Silver Spring, MD 20906

Swift, Everett R.
NSWC
Dahlgren Lab K11
Dahlgren, VA 22448

Sydnor, Richard L.
Jet Propulsion Lab
4800 Oak Grove
Pasadena, CA 91103

Tennant, Douglas M.
USAF/SAMSO
13020 Kornblum #11
Hawthrone, CA 90250

Tetu, Michel
Laval University
Elect. Eng. Dept.
Ste–Foy, Quebec, Canada

Tomlin, James, W.
Hughes Aircraft Co.
P.O. Box N
Aurora, CO 80041

Toms, Jimmie L.
Austron, Inc.
701 E. Gude Drive
Rockville, MD 20850

Underwood, P. J.
Johns Hopkins University APL
Johns Hopkins Road
Laurel, MD 20810

Van Groos, John L.
True Time Inst.
3243 Santa Rosa Avenue
Santa Rosa, CA 95401

Vanier, Prof. Jacques
Universite Laval
Quebec, Canada G1K–7P4

Venne, Pierre
Radio–Canada, Rm. #B29–1
1400 Dorchester East
Montreal, Quebec, Canada H2L–2M2

Vetter, J. R.
Applied Physics Lab
Johns Hopkins Road
Laurel, MD 20810

Vig. John R.
Army Electrnx. R&D Cmd.
Attn: DELET-MQ
Ft. Monmouth, NJ 07703

von Ulrich, Andrew
Bendix Corporation
9250 Rt. 108
Columbia, MD 21045

Walker, E. S.
Hughes Aircraft Co.
P.O. Box N
Aurora, CO 80041

Walker, William C.
Pan Am AFETR
Bldg. 989 MU 840
Patrick AFB, FL 32925

Wang, Harry T. M.
Hughes Res. Labs
3011 Malibu Canyon Road
Malibu, CA 90265

Wathen, William D., Jr.
NESEA, Code 024
St. Inigoes, MD 20684

Watling, Keith
British Embassy
3100 Massachusetts Ave., N.W.
Washington, DC 20008

Ward, Samuel C.
Jet Propulsion Lab
4800 Oak Grove Drive
Pasadena, CA 91103

Wardrip, S. C.
NASA/GSFC
Code 814
Greenbelt, MD 20771

Warren, Hugh E.
NRL
2946 Waterford Court
Vienna, VA 22180

Weidemann, Werner
Efratom California
18851 Bardeen Avenue
Irvine, CA 92646

Westerhout, Gart
U.S. Naval Observatory
Washington, DC 20390

White, Joe
Naval Research Lab, Code 7962
4555 Overlook Avenue
Washington, DC 20375

Whitworth, Gary G.
Applied Physics Lab
Johns Hopkins Road
Laurel, MD 20810

Williams, Lindsey W.
Quest Research Corp.
6845 Elm Street, Ste. 407
McLean, VA 22101

Williard, Merwin W.
Harris Corp.
P.O. Box 37
Melbourne, FL 32901

Willison, Ralph
Applied Physics Lab
9204 Crosby Road
Silver Spring, MD 20910

Wilson, Warren L.
707 Spindrift Drive
San Jose, CA 95134

Winkler, Gernot M. R.
U.S. Naval Observatory
34th and Mass. Ave., N.W.
Washington DC 20390

Woods, Bill H.
CSC
6565 Arlington Blvd.
Falls Church, VA 22046

Wolf, Stuart A.
NRL
4555 Overlook Avenue
Washington, DC 20375

Woods, Ralph A.
Electroscpace Systems
P.O. Box 1359
Richardson, TX 75080

Zich, Miles M.
NAVELEX
Washington, DC 20360

Ziegler, Robert E.
DMAHTC
6500 Brooks Lane
Washington, DC 20315